

**FIGURE 1**

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATG  
 AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAACAGAAAACCTGTTAGAAATGTGGTGGT  
 TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT  
 TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG  
 TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG  
 CTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAAA  
 TTAAACAAGGCTGGCCTTGTAAGTGGAACTAGAGTTGTTTAGGACTTCTATTGTGGCAAACCTT  
 CCAGAAAACAACCCTTTTTGTGTCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT  
 TATATATGTTTGTTCAGACCATCCTTTCCTACCAAATGCAGCCCAAAATCCATGGCAAACAAGTC  
 TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTTAGCATGCTGACTTGCTC  
 ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG  
 ACAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTT  
 GGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA  
 TGGATTAACCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA  
 GAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG  
 TTCACAGAAGTTGCTTATTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACT  
 GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCAT  
 CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACT  
 ATG

## **FIGURE 2**

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIIFSITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV  
LCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG  
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVIWCGVSALMLTCSVVLHSGNFGTDLEQKLHW  
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR  
LLSRDI

**Important features:**

**Type II transmembrane domain:**

amino acids 13-33

**Other Transmembrane domains:**

amino acids 54-73, 94-113, 160-180, 122-141

**N-myristoylation sites.**

amino acids 57-63, 95-101, 99-105, 124-130, 183-189



## FIGURE 4

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV  
 LNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTfMYFFWKLGDp  
 FPILSPKHGILSIEQLISRVGvIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERRLLQ  
 TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDALeELSRQ  
 LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDpVTRGIEI  
 TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLLAQIMGMY  
 FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQApEKQMAP

### Important features:

#### Signal peptide:

amino acids 1-23

#### Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,  
 425-444

#### N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

#### Eukaryotic cobalamin-binding proteins

amino acids 151-160

## FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC  
 AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT  
 GTTTCCTTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT  
 TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTTCCATCCAGGT  
 GTCATGCAGAATTATGGGGATCACCTTTGTGAGCAAAAAGGCGAACCAGCAGCTGAATTCACAG  
 AAGCTAAGGAGGCTGTAGGCTGCTGGGACTAAGTTTGGCCGCAAGGACCAAGTTGAAACAGCC  
 TTGAAAGCTAGCTTTGAACTTGCAGCTATGGCTGGGTGGAGATGGATTCGTGGTTCATCTCTAG  
 GATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGTGGGTGTCCTGATTTGGAAGGTTCCAGTGA  
 GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAACTCGTGCATTCCAGAA  
 ATTATCACCACCAAAGATCCCATATTCAACACTCAAACCTGCAACACAAACACAGAATTTATTGT  
 CAGTGACAGTACCTACTCGGTGGCATCCCCCTTACTCTACAATACCTGCCCCCTACTACTCTCTC  
 CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATTGATTTGTGTACAGAAGTTTTTATG  
 GAACTAGCACCATGTCTACAGAACTGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGA  
 AGCTGCTGGGTTTGGAGGTGTCCCCACGGCTCTGCTAGTGTCTGCTCTCCTCTCTTTGGTGTG  
 CAGCTGGTCTTTGGATTTTGCTATGTCAAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAAT  
 CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCATA  
 TGAGGAATCAAAGAAACTGATAAAACCCAGAAAGTCCAAGAGTCCAAGCAAAACTACCGTGC  
 GATGCCCTGGAAGCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTTCTTT  
 CATGCTCCTTACCCTGCCCCAGCTGGGGAAATCAAAAGGGCCAAAGAACCAGAAAGAAAGTCCA  
 CCCTTGGTTCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAAT  
 GCCCTTCTCCTTATTGTAACCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCCT  
 TTCTAGCCTGGCTATGTCTTAATAATATCCCACTGGGAGAAAGGAGTTTTGCAAGTGAAGGAC  
 CTAACACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGGTGGGTG  
 AAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGAGCTCAGACCCCTTTCTTCA  
 GCTCTGAAAGAGAAACACGTATCCACCTGACATGTCTTCTGAGCCCGGTAAAGAGCAAAAGAAAT  
 GGCAGAAAAGTTTAGCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAAA  
 GCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACTGTGACAGGGACTGTAAAC  
 ACAGACAGGGTCAAAGTGTCTTCTCTGAACACATTGAGTTGGAATCACTGTTTAGAACACACACA  
 CTTACTTTTTCTGGTCTCTACCCTGCTGATATTTCTCTAGGAAATATACTTTTACAAGTAACA  
 AAAATAAAACTCTTATAAATTTCTATTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT  
 ACTCAGTAATTTGTTTAAAAAGTAATAAAATTCACAAACATTTGCTGAATAGCTACTATATGTC  
 AAGTGTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATTGCACATAGTAG  
 AACGCTATCTGGGAAGCTATTTTTTTCAGTTTGTATTTCTAGCTTATCTACTTCCAACTAAT  
 TTTTATTTTGTGCTGAGACTAATCTTATTCATTTTCTCTAATATGGCAACCATTATAACCTTAAT  
 TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAGCACATTTTAAAGTGCC  
 ATTAACAAATGTATCACTAGCCCTCCTTTTCCAACAAGAGGACTGAGAGATGCAGAAATATT  
 TGTGACAAAAAATTAAAGCATTTAGAAAACCTT

**FIGURE 6**

MARCFSLVLLLSIWTTTRLLVQGSRLAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG  
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYN  
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPPAPASTSIPR  
RKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK  
RYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPESKSPSKTTVRCLEAEV

**Signal sequence:**

amino acids 1-16

**Transmembrane domain:**

amino acids 235-254

**N-glycosylation site.**

amino acids 53-57, 130-134, 289-293

**Casein kinase II phosphorylation site.**

amino acids 145-149, 214-218

**Tyrosine kinase phosphorylation site.**

amino acids 79-88

**N-myristoylation site.**

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

## FIGURE 7

CGCCGCGCTCCCGCACCCGCGGGCCCGCCACCGCGCGCTCCCGCATCTGCACCCGCGAGCCCGGC  
GGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCCGGGGCGG  
CGGCTGCGGGCGCAGAGCGGAGATGAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGCTGGCGG  
CGGCGGTCCCCACGGCCCCCGCGCCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCCG  
GCTCTCAGCTACCCGCGAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAACCTGAT  
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCGAGAAGAAGCTGCTGCTA  
AAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCAGCTATCACAATGAGACCAACACAGAC  
ACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAACAGAC  
TGGACAAATGGTCTTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC  
ACGAGTGCATCATCGACGAGGACTGTGGGGCCAGCATGTACTGCCAGTTTGGCAGCTTCCAGTAC  
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGCTGTGGAGACCA  
GCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACCAGGGGCGAGCAATGGGACCATCTGTGACA  
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCCAGAGAGGCTGCTGTTCCCTGTGTGC  
ACACCCCTGCCCGTGGAGGGCGAGCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCAC  
CTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGC  
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGG  
GAGATCCTGCTGCCCAGAGAGGTCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG  
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG  
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA  
TAGAAATAGCTAATTTATTTCCCGAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTA  
CATCTTCTTCCCAGTAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTTCAGCT  
CCCCCAGGCTGTTCTCCAGGCTTACAGTCTGGTGTCTGGGAGAGTCAGGCAGGGTTAACTGCA  
GGAGCAGTTTGGCACCCTGTCCAGATTATTGGCTGCTTTGCCCTTACCAGTTGGCAGACAGCCG  
TTTGTTCACATGGCTTTGATAATTGTTTGGAGGGAGGAGATGGAAACAATGTGGAGTCTCCCTC  
TGATTGGTTTTGGGGAATGTGGAGAAGAGTGCCCTGCTTTGCAAAACATCAACCTGGCAAAAATG  
CAACAAATGAATTTTCCACGCAGTTCTTCCATGGGCATAGGTAAGCTGTGCCTTACAGTGTGTC  
AGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTTATTCATCCAGCAGTGTGCTCAGCTCC  
TACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGG  
AGGGGGTCAATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC  
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTTCATCTGGTTGTGACTCTAAGCTCAGTGCTCT  
CTCCACTACCCACACCAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGAT  
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAACCTAATTCACATCCCTCTAAAAGTAAA  
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGAT  
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACCTTGAAGGTATATGACTGAGCGTAGCA  
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC  
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC  
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACCTTACGCCACTCCACAAATGATG  
TTTTCAGGTGTGATGACTGTTGCCACCATGTATTCATCCAGAGTCTTAAAGTTTAAAGTTGCA  
CATGATTGTATAAGCATGCTTTCTTTGAGTTTAAATATATGTATAAACATAAGTTGCATTTAGAA  
ATCAAGCATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 8**

MQRLGATLLCLLLAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL  
RSAVEEMEAEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGMVFSE  
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRLCTRDSECCGDQLCVWGHC  
TKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRLDLITWELEPDG  
ALDRPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE  
RSLTEEMALGEPAAAAAALLGGEI

**Signal sequence:**

amino acids 1-19

**N-glycosylation site.**

amino acids 96-100, 106-110, 121-125, 204-208

**Casein kinase II phosphorylation site.**

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,  
327-331

**N-myristoylation site.**

amino acids 202-208, 217-223

**Amidation site.**

amino acids 140-144



## FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC  
GGCCACCTTGTGAACCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCAAAG  
GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCGGGGCTCTTCTGGACCCTT  
AACTGGGTACTGGCCCTGGGCAATGCGTCTCGCTGGAGCCTTTGCCTCCTTCTACTGGGCCTT  
CCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC  
ACACTGGGTCAATTGGCATTGGAGCCCTCATCTGACCCTTGTGCAGATAGCCCGGGTCATCTTG  
GAGTATATTGACCACAAGCTCAGAGGAGTGACAGAACCTGTAGCCCGCTGCATCATGTGCTGTTT  
CAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCCTAAACCGCAATGCATACATCATGA  
TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAAC  
ATTGTCAGGGTGGTTCCTCGTGGACAAAAGTCACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT  
GGTCGGAGGCGTGGGGTCTGTCTTCTTTTTTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAG  
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCCTGGGGGCTAT  
GTCATCGCCAGCGGCTTCTTCAGCGTTTTCGGCATGTGTGTGGACACGCTCTTCTCTGCTTCTCT  
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA  
AGATTCTGGGCAAGAAGAACGAGGCGCCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG  
CCCTGATCCAGGACTGCACCCACCCCAACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGT  
CTCCATTTTGTGGTAAAAAAGGTTTTAGGCCAGGCGCGTGGCTCACGCCTGTAATCCAACACT  
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG  
AAACCTCCGTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCA  
GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGA  
GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAA  
AAGATTTTATTAAAGATATTTTGTTAACTC

**FIGURE 10**

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGFWTL  
 NWVLALGQCVLAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL  
 EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN  
 IVRVVLDKVTDLLLFFGKLLVVGGVVLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY  
 VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

**Important features:****Transmembrane domains:**

amino acids 57-80 (type II), 110-126, 215-231, 254-274

**N-glycosylation sites.**

amino acids 16-20, 27-31, 289-293

**Hypothetical YBR002c family proteins.**

amino acids 276-288

**Ammonium transporters proteins.**

amino acids 204-231

**N-myristoylation sites.**

amino acids 60-66, 78-84

**Amidation site.**

amino acids 306-310



## **FIGURE 12**

MGACLGACSLSCASCLCGSAPCILSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL  
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFTLLMLCVSSSRDPRAAIQ  
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSWNQRWLGKAE  
ECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVVSIAAVLPKV  
QDAQPNSSLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI  
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGVITYSY  
SFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVVKICASWAGLLLYLWTLVAPLLLRNRD  
FS

**Signal sequence:**

amino acids 1-20

**Transmembrane domains:**

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,  
272-283, 324-340, 391-406, 428-444

**FIGURE 13**

CGGGCCAGCCTGGGGCGGCCGCCAGGAACCAACCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA  
 GGTGAAAAAGACTCCTGTAACCTCCTCCAGGATGAACCACTGCCAGAAGACATGGAGAACG  
 CTCTACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA  
 CTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC  
 TTTCTGTTTGTGTCACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG  
 TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT  
 TTTGATATATTTCTTCTGGCAGTTTTTCGATTTAAAGTGTTAATACTTGCATATGCTGTGTGCAG  
 ACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAG  
 TGATCCTTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTCATTATC  
 CTTGCCTGGATTGAGACGTGGTTCCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA  
 CAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG  
 GTCAGTTTTATTCCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT  
 GAGAAACCACTTTTAGAACTATGAGTACTACTTTTTGTTAAATGTGAAAAACCTCACAGAAAGTC  
 ATCGAGGCAAAAAGAGGCAGGCAGTGGAGTCTCCCTGTGACAGTAAAGTTGAAATGGTGACGTC  
 CACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC  
 CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTCATCCTCTCT  
 TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTGTGCTGTATTCTAATC  
 AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTTATTTCATTCAC  
 AGAATGGAATTTTTTTGTTTCATGTCTCAGATTTATTTGTATTTCTTTTTTAACACTCTACATT  
 TCCCTTGTTTTTTAACTCATGCACATGTGCTCTTTGTACAGTTTTAAAAAGTGAATAAAATCTG  
 ACATGTCAATGTGGCTAGTTTTATTTTCTTGTTTTGCATTATGTGTATGGCCTGAAGTGTGGGA  
 CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAATGTCACCAGACATTTGTATTATTT  
 TTATCATGAAATCATGTTTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGC  
 ACAAATGACTTAAACCATTATATCATGTTTCCTTTGCGTTCAGCCAATTTCAATTTAAATGAA  
 CTAAATTAATAA

**FIGURE 14**

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF  
VTLLWIIELNVNGGIENTLEKEVMQYDYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT  
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLQPQAEENRLLIVQDASER  
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLEL

**Important features of the protein:****Signal peptide:**

amino acids 1-20

**Transmembrane domains:**

amino acids 54-72, 100-118, 130-144, 146-166

**N-myristoylation sites.**

amino acids 14-20, 78-84, 79-85, 202-208, 217-223



## FIGURE 16

MCSRVP LLLPL LLL L L L L L A L G P G V Q G C P S G C Q C S Q P Q T V F C T A R Q G T T V P R D V P P D T V G L Y V F E N G I T  
 M L D A G S F A G L P G L Q L L D L S Q N Q I A S L P S G V F Q P L A N L S N L D L T A N R L H E I T N E T F R G L R R L E R L Y  
 L G K N R I R H I Q P G A F D T L D R L L E L K L Q D N E L R A L P P L R L P R L L L L D L S H N S L L A L E P G I L D T A N V E  
 A L R L A G L G L Q Q L D E G L F S R L R N L H D L D V S D N Q L E R V P P V I R G L R G L T R L R L A G N T R I A Q L R P E D L  
 A G L A A L Q E L D V S N L S L Q A L P G D L S G L F P R L R L L A A A R N P F N C V C P L S W F G P W V R E S H V T L A S P E E  
 T R C H F P P K N A G R L L L E L D Y A D F G C P A T T T T A T V P T T R P V V R E P T A L S S S L A P T W L S P T A P A T E A P  
 S P P S T A P P T V G P V P Q P Q D C P P S T C L N G G T C H L G T R H H L A C L C P E G F T G L Y C E S Q M G Q G T R P S P T P  
 V T P R P P R S L T L G I E P V S P T S L R V G L Q R Y L Q G S S V Q L R S L R L T Y R N L S G P D K R L V T L R L P A S L A E Y  
 T V T Q L R P N A T Y S V C V M P L G P G R V P E G E E A C G E A H T P P A V H S N H A P V T Q A R E G N L P L L I A P A L A A V  
 L L A A L A A V G A A Y C V R R G R A M A A A A Q D K G Q V G P G A G P L E L E G V K V P L E P G P K A T E G G G E A L P S G S E  
 C E V P L M G F P G P G L Q S P L H A K P Y I

### Important features:

#### Signal peptide:

amino acids 1-23

#### Transmembrane domain:

amino acids 579-599

#### EGF-like domain cysteine pattern signature.

amino acids 430-442

#### Leucine zipper pattern.

amino acids 197-219, 269-291

#### N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

#### Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

#### N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,  
 594-600, 640-646



**FIGURE 17**

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCATGCG  
GGTCCGATAGGGCTGACGCTGCTGCTGTGTGCGGTGCTGCTGAGCTTGGCCTCGGCGTCCTCGG  
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA  
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTTCTTGATTGAGAAGAATCTGA  
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGTCACAG  
AAGATATCAGCTTTCTAGAGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTA  
CGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCT  
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTG  
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCT  
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACCTGGAATGAAAATCCTTAATGGAAG  
CAATAAGAAAAGCCAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA  
CCAAAGCCCTGGAGAGAGTGTCTATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATCCAG  
GCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGCTCTTGG  
CTTTCTGTATGCCTCTGGACTTGGTGTTAATTCAAGTCAGGCAAAGGCTCTTGATATTATACAT  
TTGGAGCTCTTGGGGCAATCTAATAGCCCACATGGTTTTGGTAAGTAGACTTTTAGTGGAAGGCT  
AATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACCTTTTTCAGCTTTTCATGATC  
CAGATTTGCTTGTTATTAAGACCAAATATTCAGTTGAACTTCCTTCAAATCTTGTTAATGGATAT  
AACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACAATTTTCTTTAAATGATTAG  
TTTGGCTGATTGCCCCATAAAAAGAGAGATCTGATAAATGGCTCTTTTAAATTTTCTCTGAGTTG  
GAATTGTCAGAATCATTTTTTACATTAGATTATCATAATTTTAAAAATTTTCTTTAGTTTTTCA  
AAATTTTGTAAATGGTGGCTATAGAAAAACAACATGAAATATTATACAATATTTTGCAACAATGC  
CCTAAGAATTGTTAAATTCATGGAGTTATTTGTGCAGAATGACTCCAGAGAGCTCTACTTTCTG  
TTTTTTACTTTTCATGATTGGCTGTCTTCCATTTATTCTGGTCATTTATTGCTAGTGACACTGT  
GCCTGCTTCCAGTAGTCTCATTTTCCCTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGG  
AAGATTAACTCATTTTAAATAAATTATGTCTAAGATTAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAA

## **FIGURE 18**

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLSKTTLTSDSVKDHTTAGRVVAGQIFLDSESESEL  
 ESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHGEPCHFPLFLDK  
 EYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEAAKRRQMQEAEEMMYQTGMKILNGSNKKSQKR  
 EAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMFELTEEGSPKGQTALGFLYASGLGVN  
 SSQAKALVYYTTFGALGGNLIAMVLVSRL

**Important features:**

**Signal peptide:**

amino acids 1-21

**N-glycosylation sites.**

amino acids 195-199, 217-221, 272-276

**Tyrosine kinase phosphorylation site.**

amino acids 220-228

**N-myristoylation sites.**

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

**Glycosaminoglycan attachment site.**

amino acids 267-271

**Microbodies C-terminal targeting signal.**

amino acids 299-303

**Type II fibronectin collagen-binding domain protein.**

amino acids 127-169

**Fructose-bisphosphate aldolase class-II protein.**

amino acids 101-119

## FIGURE 19

AATTCAGATTTTAAGCCCATCTGCAGTGGAATTCATGAAC TAGCAAGAGGACACCATCTTCTT  
GTATTATACAAGAAAGGAGTGACCTATCACACACAGGGGGAAAAATGCTCTTTTGGGTGCTAGG  
CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAAC TAAAGATTGAAGACATCACTG  
ATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAAC TTTTGAT  
AAAAAGGGATTTTCATGTAATCGCTGCCGTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA  
AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG  
CCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCAATAATGCTGGTGTT  
CCCGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA  
CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAG  
TTATTAATGTCTCCAGTGTGGAGGTGCCTTGCAATCGTTGGAGGGGGCTATACTCCATCCAAA  
TATGCAGTGGAAGGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGTCTC  
ATGCATTGAACCAGGATTGTTCAAACAAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC  
TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAACAATATGGAGAAGGTTACATTGAAAAA  
AGTCTAGACAAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG  
CATGGACCACGCTCTAACAAGTCTCTTCCCTAAGACTCATTATGCCGCTGGAAGATGCCAAAA  
TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAA  
GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCTCCAGGCTATGA  
AATTGGCCGATTTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACT  
CATTTAGATCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCAGGGT  
CCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCT  
GTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATGA  
TCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCCAGCATTTACAGTAACTTGGAATGTTAAGT  
ATCATCTCTTATCTAAATATTAAAAGATAAGTCAACCCAAAAA AAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAA

**FIGURE 20**

MLFWVLGLLILCGFLWTRKGKLIKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG  
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNVQGEKGLWGLINNAGVPGVLAPTDWLTLEDY  
REPIEVNLFGLISVTNLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK  
AFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKGNKSYVNMD  
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAE LANPKAV

**Important features of the protein:****Signal peptide:**

amino acids 1-17

**Transmembrane domain:**

amino acids 136-152

**N-glycosylation sites.**

amino acids 161-163, 187-190 and 253-256

**Glycosaminoglycan attachment site.**

amino acids 39-42

**N-myristoylation sites.**

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

## FIGURE 21

CTGAGGCGGCGGTAGCAATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGCG  
 CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGTAAAA  
 GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA  
 CATTAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAGGCGAAGTAAATG  
 AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCCGT  
 CGTCATTAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACTTGCAGGAGCATT  
 TTCAAACCAAGACCTTGTTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTCTACTC  
 ATCGACTGGAACATTCTTATATAAACCTCAAAAAGGACTTTTTTCACAGGGTACCTTTAGTGGTT  
 GCCAATCTGGGCATGCTGAACAACGGGTATAAACTGTATCAGGTTCTGTATGTCCACTGG  
 TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTGAAGAAGATGGATCCTTAAAGGAGG  
 TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAGTG  
 GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAAACGAGAAATTGA  
 GAAAAGGAGAGGAGCACAGATTAGGCAGCAAGAGAGAAGAACATCCAAAAGACCTCAGGAGA  
 ACATTTTTCTTTGTCAGGCATTACGGACCTTTTTTCCAAATCTGAATTTCTTCATTCATGTGTT  
 ATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAACACACCACCATCTCGATGTAGT  
 AGACAATCTGACCTTAATGGTAGAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC  
 AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTA  
 GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAT  
 GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTC  
 CTACATTTTGAATCCTTTTAACTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAAC  
 ATTTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTTCATTTGTTTTTACTATGTTTACC  
 TGTTTGAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAACATCA  
 GATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCACCTAAGTTGTTGAGGGGAAGGCTTACACAG  
 ACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCAGCACT  
 TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT  
 GAGACCATGTCTATTAAAAAATAAATGGAAAAGCAAGAATAGCCTTATTTTCAAATATGGA  
 GAAATTTATATGAAATTTATCTGAGTCATTAATTTCTCCTTAAGTGATACTTTTTTAGAAGTA  
 CATTATGGCTAGAGTTGCCAGATAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT  
 AAAATTTAAAAA

**FIGURE 22**

MEGESTSAVLSGFVLGALAFOHLNTDSDEGFLGGEVKGEAKNSITDSQMDDVEVVYTIDIQKYI  
 PCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL  
 VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV  
 QTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA  
 QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCYNHHLVDVVDNLTL  
 MVEHTDIPEASPASTPQIIKHKALDLDLRWQFKRSRLLDTDKRSKANTGSSNQDKASKMSSPET  
 DEEIEKMKGFGEYSRSPTE

**Important features:****Signal peptide:**

amino acids 1-19

**N-glycosylation sites.**

amino acids 75-79, 322-326

**N-myristoylation site.**

amino acids 184-154

**Growth factor and cytokines receptors family.**

amino acids 134-150

## FIGURE 23

GGCACAGCCGCGCGGGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCCCAA  
GCAGCGCGCAGCGAACGCCCGCGCCGCCACACCCCTCTGCGGTCCCCGCGGCGCCTGCCACCCCTCCCTCCTTCCCC  
GCGTCCCCGCGCTCGCCGGCCAGTCAGCTTGGCGGTTGCTGCCCCGCGAAACCCGAGGTACACAGCCGCGCCTCT  
GCTTCCCTGGGCGCGCGCGCCTCCACGCCCTCCTTCTCCCTGGCCCGGCGCCTGGCACCGGGGACCCTTGCCTGA  
CGCGAGGCCCAGCTCTACTTTTCGCCCGCGCTCTCCTCGCCTGCTGCGCTCTTCCACCAACTCCAACCTCCTTCTCCC  
TCCAGCTCCACTCGCTAGTCCCGGACTCCGCCAGCCCTCGGCCCGCTGCGGTAGCGCGCTTCCCGTCCGGTCCCAAA  
GGTGGGAACGCGTCCGCCCGCGCCGACCATGGCACGGTTGCGGTTGCCCGCGCTTCTCTGCACCTGGCAGTGCTC  
AGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAGTTGCTCGGAAGTGCAGCTCTTACGTGTCCAAAGGCTTC  
AACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCATTGAAGATCTGTCCCAGGGTTCTACCTGCTGCTCT  
CAAGAGATGGAGGAGAAGTACAGCTGCAAGTAAAGATGATTCAAAGTGTGGTCAGCGAACAGTGAATCATTG  
CAAGCTGTCTTTGCTTACGTTACAAGAAGTTTGATGAATTCTCAAAGAACTACTGAAAATGCAGAGAAATCCCTG  
AATGATATGTTGTGAAGACATATGGCCATTTATACATGCAAAATTCTGAGCTATTTAAAGATCTCTTCTAGAGTTG  
AAAGCTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATG  
TTCGCGCTGGTGAACCTCCAGTACCCTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGAGCAGCTGAAG  
CCCTTCGGAGATGTCCCTCGCAAATTGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTCTGCTCAAGGC  
TTAGCGGTTGCGGGAGATGTCTGAGCAAGGTCTCCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG  
ATGATCTACTGCTCCCACTGCCGGGCTCTCGTGACTGTGAAGCCATGTTACAATACTGCTCAAACATCATGAGAGGC  
TGTTTGGCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTATAGATGCTATGCTGATGGTGGCAGAGAGGCTA  
GAGGCTCCTTTCAACATTGAATCGGTGATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGAT  
AATAGTGTTCAGTGTCTCAGAAGGTTTCCAGGGATGTGGACCCCAAGCCCTCCAGCTGGACGAATTTCTCGT  
TCCATCTCTGAAAGTGCCTTCAGTGCTCGCTTCAGACCACATCACCCGAGGAACGCCCAACACAGCAGCTGGCACT  
AGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTTCTGGTCTCCTTCCGAGCAAC  
GTTTGCAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC  
CTGTTTGCACTGACAGGAAATGGATTAGCCAACAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAACACAGAC  
ATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGATGCATACAATGGGAACGACGTG  
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAGTGAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG  
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGCTCCTGGGGCA  
CAGGCCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCTGGTTATGCAGAGAGAGTGGAGATTAATTCTCAAACCTGAG  
AAAAAGTGTTCATCAAAAAGTTAAAAGCACCAGTTATCACTTTTCTACCATCCTAGTGACTTTGTTTTTAAATGAA  
TGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAAAGTGTGACTTTGTTTTCTCATTGATTTTGGG  
AGGAAAAGGGACTGTGCATTGAGTTGGTTCTGCTCCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA  
CTATAGTTAGTTGTGCATTTGTGATTTTATCACTCTATTATTTGTTGTATGTTTTTTCTCATTTCGTTTGTGGGTT  
TTTTTTTCCAACGTGATCTCGCCTTGTCTTACAAGCAAACAGGGTCCCTTCTGGCACGTAACATGTACGTATT  
TCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTTATTATCATGTTATCTTATTAAGAAAAAGCCCAAAAGC

## FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST  
CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH  
LYMQNSELFKDLFVELKRYVVGNNLEEMLNDFWARLLERMFLVNSQYHFTDEYLECVSKYTE  
QLKPFQDVPRKLLQVTRAFVAARTFAQGLAVAGDVVSKVSNNPTAQCTHALLKMIYCSHCRGL  
VTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN  
MQDNSVQVSQKVFQGC GPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK  
EKLKQAKKFWSSLPSNVCNNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNEVQVDT  
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGGCEYQCCPSEFDYNATDHAG  
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

**Important features:**

**Signal peptide:**

amino acids 1-22

**ATP/GTP-binding site motif A (P-loop).**

amino acids 515-524

**N-glycosylation site.**

amino acids 514-518

**Glycosaminoglycan attachment sites.**

amino acids 494-498, 498-502

**N-myristoylation sites.**

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

**Glypicans proteins.**

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506



## FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC  
 CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT  
 TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA  
 TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGACCGAGGC  
 CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCTGAG  
 AGCCCCGAGAAGAAAATTTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCCCTGTGATCATT  
 TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA  
 GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG  
 AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC  
 TTCTTCTCCACCTCACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA  
 TGTTTTTCAAGATCATTTTGTGTTGTTGCTCTCTCTAGTGCTTCTTCTCTCGTCAGTCTTAGCCT  
 GTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCT  
 AGCTAGTGTCATTTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT  
 AAATGTCAAAAAAAAAAAAAAAAAA

## **FIGURE 26**

MKVLISLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM  
TVSGLPKKQCPCDHFKNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

**Important features:**

**Signal peptide:**

amino acids 1-22

**N-myristoylation sites.**

amino acids 27-33, 46-52

**FIGURE 27**

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG  
 AGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTACCCTGCC  
 CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCCCCTACCTGA  
 TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG  
 ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC  
 CAACTTTTCACTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA  
 AGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT  
 GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGTGTG  
 CTCTGTGCAGAGCCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC  
 TCTTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCCTCATGTGGCAGCAAGTTTTT  
 GAGCCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA  
 GAACGCCCAGTTCTCCGAAATCCAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG  
 GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACTCATTGTCTCC  
 TTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTA  
GCAGAATGAGAGAAGACATTATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC  
 AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG  
 AAACACTAGGACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTC  
 CCAATGTTGTCCCTTTCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCTCTGAGGCTACAC  
 CCATGCGTCTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC  
 CCTCTCTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGG  
 ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTCTCAAATATTTTTTAATAAATAGACGAA  
 ACCACG

**FIGURE 28**

MDILVPLLQLLVLLLTPLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSSQIKGL  
TGASGKVALLELGCGTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM  
RQLADGSMDVVVCTLVLCSVQSPRKVLQEVRRVLRPGGVLEFFWEHVAEPYGSWAFMWQQVFEPTW  
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFSSKALICSFPSL  
QLEQATHQPIYLPLRGT

**Important features:****Signal peptide:**

amino acids 1-23

**Leucine zipper pattern.**

amino acids 10-32

**N-myristoylation sites.**

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

## FIGURE 29

CAATGTTTGCCTATCCACCTCCCCAAGCCCCTTTACCTATGCTGCTGCTAACGCTGCTGCTGCT  
 GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCAGAAAGTCTCTTCTG  
 CCACTGACGCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTGCCTCAT  
 CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGCTA  
AGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTTAAA  
 GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTTCATTAAGAAATT  
 GTTCCCTTCCCCTGTGTTCAATGTTTGTAAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC  
 AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

### **FIGURE 30**

MLLLTLLLLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQPRG  
EGEKVGDG

**Important features:**

**Signal peptide:**

amino acids 1-15

**Growth factor and cytokines receptors family:**

amino acids 3-18

# **FIGURE 31**

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCAGTGTGTCCCAGGCTACCAAGTT  
CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT  
ATTTGCATCTGTTTTGATAAATGATGTTGACACCTCCACCGAATCTAAGTGGAATCATGTCGG  
GAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGATGTTTACC  
TTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTATTGGTTATTTTGGGATTGTTGTT  
TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA  
CAGAAAGGGGAAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG  
CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCAC  
AAATAAAGCCATCAGCAGTGCTCCCTTCTGCTGTTCCAGCCACTGTGGACATTTGCCATCCTCA  
TTTTCTTCTGGGTCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAACCTGCAGGAGCTGCCCAGGTT  
ATGGAAGGCGGCCAAGTGAATATAAGCCCTTTTCGGGCATTTCGGTACATGTGGTCGTACCATTT  
AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGGCTGCCAGCAATGACTATAGCTGGGGCAG  
TGGTTACTTGTTATTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTC  
TCCATTCTCTTCTTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAG  
GATTCCGAGAATCATTGTCATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCAATTGT  
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC  
AACCAGAATGCATATACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC  
ATTCAAAATCTTGTCCAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAA  
TTTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACAC  
AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATGGTAGCTTTTTTGCCTACTTAGTAGC  
CCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCTGTGTTTTGCTGTTGATC  
TGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA  
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAAGGAATGAGGA  
GGGAACAGAACTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCTGGAAAACATT  
TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTT  
TTAAAGACCTAATAAACCTATTCTTCCTCAAAA

## FIGURE 32

MSGRDTILGLCILALALSLAMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYTNDLSIE  
 LDTERENMKCVLGFAIVSTGITAVLLVLI FVLKRRIKLTVELFQITNKAISSAPFLLFQPLWTFA  
 ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQOMTIA  
 GAVVTCYFNRSKNDDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMYMQNALKEQQHG  
 ALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD  
 FIIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA  
 VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

### Important features:

#### Signal peptide:

amino acids 1-20

#### Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

#### N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

#### N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348



**FIGURE 33**

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTTGGACCTCTCCCTGTTTCTTCCTTAGA  
 ATAATTTGTATGGGATTTGTGATGCAGGAAAGCCTAAGGGAAAAAGAATATTCTGTGTGGT  
 GAAAATTTTTTGAAAAAAATTCGCTTCTTCAAACAAGGGTGTCTATCTGATATTTATGAGGAC  
 TGTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTTCTTGTGTTTGTGGTGAAGTAC  
 ATTCAAACAAAGAACGGCAAAGAAGATTAAAAGGCCAAGTTCAGTGTGCCTCAGATCAACTGC  
 GATGTCAAAGCCGGAAGATCATCGATCTGAGTTCATTGTGAAATGTCCAGCAGGATGCCAAGA  
 CCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGTGGCGCTGCCG  
 TACACAGTGGTGTGCTTGATAATTCAGGAGGGAAAAATACTTGTTCGGAAGGTTGCTGGACAGTCT  
 GGTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTTATCCCTACCACGATGGAGAGAATCCTT  
 TATCGTCTTAGAAAGTAAACCCAAAAAGGGTGTAACTTACCCTACCTATCAGCTCTTACATACTCATCAT  
 CGAAAAGTCCAGCTGCCCAAGCAGGTGAGACCACAAAAGCCTATCAGAGGCCACCTATTCCAGGG  
 ACAACTGCACAGCCGGTCACTCTGATGCAGCTTCTGGCTGTCACTGTAGCTGTGGCCACCCCCAC  
 CACCTTGCCAAAGGCCATCCCCCTTCTGCTGCTTCTACCACCAGCATCCCCAGACCACAATCAGTGG  
 GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAAACAGGCC  
 AGAGCTGATCCAGGTATCCAAAGGCAAGATCCTTCAGGAGCTGCCTTCCAGAAACCTGTTGGAGC  
 GGATGTGAGCTGGGACTTGTTCAAAAGAAGAATTGAGCACACAGTCTTGGAGCCAGTATCCC  
 TGGGAGATCCAACTGCAAAATTGACTTGTGCTTTTTAATTGATGGGAGCACCAGCATGGCAAA  
 CGGCGATTCCGAATCCAGAAGCAGCTCCTGGCTGATGTTGCCAAGCTCTTGACATTGGCCCTGC  
 CGGTCCACTGATGGGTGTTGTCCAGTATGGAGACAACCCTGCTACTCACTTTAACCTCAAGACAC  
 ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTTCTAAT  
 GTAGGTTCGGGCCATCTCCTTGTGACCAGAAGTCTTTTCCAAAGCCAATGGAACAGAAAGCGG  
 GGCTCCCAATGTGGTGGTGGTGTGATGGTGGATGGCTGGCCACGGACAAAAGTGGAGGAGGCTTCAA  
 GACTTGGGAGAGATCAGGAATCAACATTTTCTTCATCACCATTGAAGGTGCTGCTGAAAATGAG  
 AAGCAGTATGTGGTGGAGCCCACTTTGCAAAACAAGGCCGTGTGCAGAACAAACGGCTTCTACTC  
 GCTCCACGTGCAGAGCTGGTTTGGCCTCCACAAGACCTGCAGCCTCTGGTGAAGCGGGTCTGCG  
 AACTGACCGCTGGCCTGCAGCAAGACCTGCTTGAACCTGGCTGACATTGGCTTCGTCATCGAC  
 GGCTCCAGCAGTGTGGGACGGGCAACTTCCGCACCGTCTCCTCAGTTTGTGACCAACCTCACCAA  
 AGAGTTTGAGATTTCCGACACGGACACGCGCATCGGGGCCGTGCAGTACACCTACGAACAGCGGC  
 TGGAGTTTGGGTTGACAAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGGC  
 TACTGGAGTGGTGGCACCAGCAGCGGGGCTGCCATCACTTCGCCCTGGAGCAGCTCTTCAAGAA  
 GTCCAAGCCCAACAAGAGGAAGTTAATGATCCTCATCACCGACGGGAGGTCCTACGACGACGTCC  
 GGATCCCGCATGGCTGCCCATCTGAAGGGAGTGATCACCTATGCGATAGGCGTTGCTTGGCT  
 GCCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCCGCCAGAGACCACTCCTTCTTGTGGACGA  
 GTTTGACAACCTCCATCAGTATGTCCCAGGATCATCCAGAACATTGTACAGAGTTCAACTCAC  
 AGCCTCGGAACGTAATTCAGAGCAGGCAGAGCACCAGCAAGTGTGCTTTACTAACTGACGTGTT  
 GGACCACCCACCGCTTAATGGGGCACGCACGGTGCATCAAGTCTTGGGCAGGGCATGGAGAAAC  
 AAATGTCTTGTATTATTCTTTGCCATCATGCTTTTTTCATATTCCAAAACCTGGAGTTACAAAGA  
 TGATCACAAACGTATAGAATGAGCCAAAAGGCTACATCATGTTGAGGGTGTGGAGATTTTACAT  
 TTTGACAATTGTTTTCAAATAAATGTTGGAATACAGTGCAGCCCTTACGACAGGCTTACGTAG  
 AGCTTTTGTGAGATTTTAAAGTTGTTATTCTGATTGAACTCTGTAACCCTCAGCAAGTTTCAT  
 TTTTGTGATGACAATGTAGGAATTGCTGAATTAATGTTTAGAAGGATGAAAAATAAAAA  
 AA  
 AAG

## FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG  
 CQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYNGVQSLSLPRWR  
 ESFIVLESKPKKGVITYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA  
 TPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTSSQNRPRADPGIQRQDPSGAAFQKP  
 VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI  
 GPAGPLMGVVQYGDNPATHFNLKTHHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN  
 RSGAPNVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG  
 FYSLHVQSWFGLHKTQLPLVKRVCDTDLACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN  
 LTKEFEISDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALEQL  
 FKKS KPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAWAAQEELEVIATHPARHSFF  
 VDEFDNLHQYVPRIIQNICTEFNSQPRN

### Important features:

#### Signal peptide:

amino acids 1-26

#### Transmembrane domain:

amino acids 181-200

#### N-glycosylation sites.

amino acids 390-394, 520-524

#### N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,  
 431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

#### Amidation site.

amino acids 304-308

**FIGURE 35**

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC  
 CAAACCATGTCTTTTTTCTGTTTTTCAGAGTAGTTCCACACAGATCTGAGTGTTTTAATTAAGCATGGAAT  
 ACAGAAAACAACAAAACTTAAGCTTTAATTCATCTGGAATCCACAGTTTTCTTAGCTCCCTGGACCC  
 GGTGACCTGTTGGCTCTTCCGCTGGCTGCTCTATCACGTGGTGCTCTCCGACTACTACCCCGAGTGTA  
 AAGAACCTTCGGCTCGCGTGCTTCTGAGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGTCTTCCGAGTA  
 GGATGTCACTGAGATCCCTCAAATGGAGCCTCCTGCTGCTGCTCACTCCTGAGTTTCTTTGTGATGTGGTAC  
 CTCAGCCTTCCCACTACAATGTGATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTTA  
 CAGACAAGACTTTCACCTTCACACTTCGAGAGCATTCAAAGTCTCTCATCAAAATCCATTTCTGGTCATTCT  
 TGGTGACCTCCCACCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTGAAAAAAGTCT  
 TGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAGAGGCTGAAAGGAAGACAAAATGTTGGC  
 ATTGTCTTAGAGGATGAACACCTTCTTATGGTGACATAATCCGACAAGATTTTTTAGACACATATAATA  
 ACCTGACCTTGAAACCATTTATGGCATTGAGGTGGGTAACTGAGTTTGGCCCAATGCCAAGTACGTAATG  
 AAGACAGACACTGATGTTTTCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGA  
 GAAGTTTTTCAGAGTTATCCTCTAATTGATAATTATCCTATAGAGGATTTTACCAAAAACCCATATTT  
 CTTACCAGGAGTATCCTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTTATATAATGTCCAGAGAT  
 TTGGTGCCAAGGATCTATGAAATGATGGGTACGTAAAACCCATCAAGTTGAAGATGTTTATGTCCGGAT  
 CTGTTTGAATTTATTAAAAGTGAACATTCATATTCAGAAGACACAAATCTTTCTTTCTATATAGAATCC  
 ATTTGGATGTCTGTCAACTGAGACGTGTGATTGCAGCCCATGGCTTTCTTCCAAGGAGATCATCACTTTT  
 TGGCAGGTCATGCTAAGGAACACCACATGCCATTATTAAGCTTCACATTCTACAAAAGCCTAGAAGGACAG  
 GATACCTTGTGGAAAGTGTTAAATAAAGTAGGTACTGTGGAAAATTCATGGGGAGGTGAGTGTGCTGGCTT  
 AACTGAACTGAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTATTAGTCAGG  
 CCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAAGAAATTAATAGG  
 ACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAAGGTGTTACTGAGTTATAAGCTCA  
 CTAGGCTGTAAAAACAAAACATGTAGAGTTTATTTATTGAACAATGTAGTCACTTGAAGGTTTTGTGTA  
 TATCTTATGTGGATTACCAATTTAAAAATATATGTAGTTCTGTGTCAAAAACTTCTTCACTGAAGTTATA  
 CTGAACAAAATTTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATT  
 ATTATTTAAATTAATCACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAGGATAG  
 TGAATCATTTCTTACATGCAACATTTCCAGTACTTAACTGATCAGTTTATTATTGATACATCACTCCA  
 TTAATGTAAAGTCATAGGTCAATTATTGCATATCAGTAATCTCTTGGACTTTGTTAAATATTTTACTGTGGT  
 AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

## **FIGURE 36**

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF  
TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWSWGVEVLTFFLLGQAEKEDKMLA  
LSLEDEHLLYGDIIRQDFLDTYNNLTLKTIMAFRWVTEFCPNAKYVMKTDVDFINTGNLVKYLE  
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYFPKVFPPYCSGLGYIMSRDLVPRIYEMMGHV  
KPIKFEDVYVGICLNLKVNIIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR  
NTTCHY

**Important features:**

**Type II transmembrane domain:**

amino acids 20-39

**N-glycosylation sites.**

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

**Glycosaminoglycan attachment site.**

amino acids 239-243

**Ly-6 / u-PAR domain proteins.**

amino acids 23-37

**N-myristoylation site.**

amino acids 271-277

**FIGURE 37**

CGCTCGGGCACCAGCCGCGGCAAGGATGGAGCTGGGTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCCTTCTCATC  
TCGTCCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAATATCATGTGTCGGGAGTGCTGTG  
AATATGATCAGATTGAGTGCGTCTGCCCCGAAAGAGGAAGTCGTGGGTATACCATCCCTTGCTGCAGGAATGAGGAGAA  
TGAGTGTGACTCCTGCGTGTATCCACCCAGGTTGTACCATCTTTGAAAAGTGAAGAGCTGCCGAAATGGCTCATGGGGGGT  
ACCTTGAGTACTTCTATGTGAAGGGGTTCTACTGTGCAGAGTGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTG  
GCCAGGTTCTGCGAGCCCCAAGGGTCAGATTTTGTGGAAAGCTATCCCTAAATGCTCACTGTGAATGGACCATTTCATGC  
TAAACCTGGGTTTGTCTATCCAACATAAGATTTGTCTGTGAGTCTGGAGTTGACTACATGTGCCAGTATGACTATGTTGAG  
GTTCTGTATGGAGACAACCGCGATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAG  
GATCCTCACTCCACGTCTCTTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTCCATGCCATTTATGAGGAGATCACAGC  
ATGCTCCTCATCCCTTGTTTCCATGACGGCACGTGCGTCTTGACAAGGCTGGATCTTACAAGTGTGCTGCTTGGCAGGC  
TATACTGGGCAGCGCTGTGAAAATCTCCTTGAAGAAAGAACTGCTCAGACCTTGGGGGCCAGTCAATGGGTACCAGAAAA  
TAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAATTTGGCACCCTGGTGTCTTTCTTTTGAACAACTCTATGT  
TCTTAGTGGAATGAGAAAAAGAACTTGCCAGCAGAATGGAGAGTGGTCAGGGAACAGCCCATCTGCATAAAAGCTTGCCGA  
GAACCAAGATTTTCAAGCCTGGTGAGAAAGGAGATTCTTCCGATGCAGGTTCAAGTCAAGGGAGACACCATACACCAGCTAT  
ACTCAGCGGCCCTTCAGCAAGCAGAACTGCAGAGTGCCCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCATGGG  
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTCACCTTCTACCGCCGCTGGGCAGCAGCAGGAGGACATGT  
CTGAGGACTGGGAAGTGGAGTGGGGGGCACCATCTGCATCCCTATCTGCGGGAAAATTGAGAACATCACTGCTCCAAAGA  
CCCCAAGGGTTGCGCTGGCCGTGGCAGGCAGCCATCTACAGGAGGACCAGCGGGGTGCATGACGGCAGCCTACACAAGGGAGC  
GTGGTTCTTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCCACTGTGTTACTGACCTGGGGAAG  
GTCACCATGATCAAGACAGCAGACTGAAAAGTTGTTTGGGGAAATTTACCGGGATGATGACCGGGATGAGAAGACCATCC  
AGAGCCTACAGATTTCTGCTATCATTTCTGCATCCCAACTATGACCCCATCTGCTTGATGCTGACATCGCCATCTGAAGCT  
CCTAGACAAGGCCGTATCAGCACCCGAGTCCAGCCCATCTGCCCTCGCTGCCAGTCGGGATCTCAGCACTTCTTCCAGGAG  
TCCCACATCACTGTGGCTGGCTGGAATGTCTGGCAGACGTGAGGAGCCCTGGCTTCAAGAACGACACACTGCGCTCTGGGG  
TGGTCAGTGTGGTGGACTCGCTGCTGTGTGAGGAGCAGCATGAGGACCATGGCATCCCACTGAGTGTCACTGATAACATGTT  
CTGTGCCAGCTGGGAACCACTGCCCTTCTGATATCTGCACTGCAGAGACAGGAGGCATCGCGGCTGTCTCTTCCCGGGA  
CGAGCATCTCCTGAGCCACGCTGGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCA  
CTGCCCTTACCAAGGTGCTGCCCTTTAAAGACTGGATTGAAAGAAATATGAAATGAACCATGCTCATGCACTCCTTGAGAAG  
TGTTTCTGTATATCCGTCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGGCTGAAGTGTGATTTGGCCTGTGAACCTGGCT  
GTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGTTGAGTAGACCTCCATTGCTGGTAGGCTGATGCCGCTCCA  
CTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGCAAGAAGTAAGTTCTTCAAGAAGACCATATACAAAACCTCTCCA  
CTCCACTGACCTGGTGGTCTTCCCCAACTTTCAAGTATACGAATGCCATCAGCTTGACCAGGGAAGATCTGGGCTTCATGAG  
GCCCTTTTGGGCTCTCAAGTTCTAGAGAGCTGCCCTGTGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCCTT  
TGTGTACATGGCCACAGTACAGTCTGGTCTCTTTCTCTCCCCATCTCTGTACACATTTAATAAAATAAGGGTTGGCTCT  
GAAC TACAAA  
AAA

## FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT  
 IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGFYCAECRAGWYGGDCMRCGQ  
 VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGOII  
 KRVCGNERPAPIQSIGSSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC  
 ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNSYVLSGNE  
 KRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPVQVSRETPLHQLYSAAFSKQKLQSAPTK  
 KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRRTCLRTGKWSGRAPSCIPICGKIENITAP  
 KTQGLRWPWQAAYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL  
 KVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR  
 DLSTSFQESHITVAGWNVLADVRS PGFKNDTLRSGVSVVDSLLCEEQHEDHGIPVSVTDNMFCA  
 SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMGLVSWSYDKTCSHRLSTAFTKVLFPKDWI  
 ERNMK

### Important features of the protein:

#### Signal peptide:

amino acids 1-23

#### EGF-like domain cysteine pattern signature.

amino acids 260-272

#### N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

#### N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,  
 474-480, 491-497, 638-644, 666-672

#### Amidation site.

amino acids 56-60

#### Serine proteases, trypsin family.

amino acids 489-506

#### CUB domain proteins profile.

amino acids 150-167

## FIGURE 39

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTTATTAACTGGCTTAATC  
TGAAGGTTCTCAGTCAAATCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGTTTGCTTAAAGGAGCTTGGCTGG  
TTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAAATGAAGGCGCTTCTGTTGC  
TGGTCTTGCCCTGGCTCAGTCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCTGTATTGAGAACTCTGTA  
AAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAGACGGCTGTGCGAGCC  
TCACAGCCACGGCTCCCTCCCCAGAGGTTTCTGCAGCTGCCACCATCTCCTTAATGACAGACGAGCTTGGCCTAGACA  
ACCTTGCTACGTGTCTCGGCAGAGGACGGGCGAGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGAACTA  
GGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAAATCGAGCTTTGAGTGTCTTCGAA  
GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCGGCGAGGAAAATTCTGAAACACCACTGCCCTGAAG  
TCTTTCAAGGTTGTACCACTGATTCCAGATGGTGAAATTACCAGCATCAAGATCAATCGAGTAGATCCAGTGAA  
GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCCATATCATTTCAACACATTTATCGTGATGGG  
TGATGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC  
ACAACTACGCTGTGCTCTCTGCGGCAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCCGCA  
GCAGGAACAATGGACAGGCCCGGATGCCTACAGACCCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCC  
CCGAGGAGCAGCTTGAATAAACTGGTGCAGGATGGATGAGCCTGGGGTTTTTCATCTTCAATGTGCTGGATGGCG  
GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTAGCCATCAATGGACATGATCTTCGATATGGCA  
GCCCAGAAAGTGCGGCTCATCTGATTACGGCCAGTGAAAGACGTGTTACCTCGTGTCTCCGCCAGGTTCCGCGAGC  
GGAGCCCTGACATCTTTCAGGAAGCGCGCTGGAACAGCAATGGCAGCTGGTCCCGAGGCGAGGGGAGAGGAAACA  
CTCCCAAGCCCTCCATCTTACAATTACTTGTCTGATGAGAAGGTGGTAAATATCCAAAAAGACCCCGGTGAATCTCTCG  
GCATGACCGTCGAGGGGAGCATCACATAGAGAATGGGATTTGCCATCTATGTCATCAGTGTGAGCCCGGAGGAG  
TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTTTGTGAATGTGGATGGGGTCGAAGTACAGAGGTGAGCC  
GGAGTGAGGCACTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGC  
CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGACTGGTCCCCATCCT  
GGGTCATGTGGCTGGAATTACCAAGGTGCTTGTATAACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGTC  
CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATCTTCTTGCTGTCAATGGTAGAAGTACATCAGGAATGA  
TACATGCTTGCTTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTGGCACTT  
TTTTATAGAATCAATGATGGGTGAGAGGAAAACAGAAAAATCACAAATAGGCTAAGAAGTTGAAACACTATATTATC  
TTGTCAAGTTTTATATTTAAAGAAAGAAATACATTGTAAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTT  
ACACCTCAGAAAAATATGATTCAAAAAAATTAACCTACTAGTTTTTTTTCAGTGTGGAGGATTTCTCATTAATCTAC  
AACATTGTTTATATTTTCTATTCAATAAAAGCCCTAAACAACTAAAATGATTGATTGTATACCCCACTGAATT  
CAAGCTGATTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTAAATTACAGCT  
AAAAATTTTTTAAATGCATTGCTGAGAAACGTTGCTTTCATCAACAAGAATAAATATTTTTTCAGAAGTTAA

## **FIGURE 40**

MKALLLLVLPWLS PANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATAPS  
 PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSFKKINR  
 ALSVLRRTKSGSAVANHADQGRESENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV  
 GGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM  
 REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVKRVDEPGVFIFNVLDGGVAYRHG  
 QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQRSPDIFQEAGWNSNGSWSPG  
 PGERSTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDR  
 IKTGDIILNVGDVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAP  
 SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG  
 RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

### **Important features:**

#### **Signal peptide:**

amino acids 1-15

#### **N-glycosylation sites.**

amino acids 108-112, 157-161, 289-293, 384-388

#### **Tyrosine kinase phosphorylation sites.**

amino acids 433-441, 492-500

#### **N-myristoylation sites.**

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,  
 467-473, 603-609



**FIGURE 41**

ACCAGGCATTGTATCTTCAGTTGTTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT  
 CTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTCACGTAATAAAAAACATGGGCTTCAACCTGACT  
 TTCCACCTTTTCCTACAAATCCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGGTTGGGTGGGC  
 CACCAGTAACACTTTCGTGGGTGCCATTCAAGAGATTCCCTAAAGCAAAGGAGTTCATGGCTAATTTCC  
 ATAAGACCCCTCATTTTGGGAAGGGAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACTT  
 GACAACTGTCTTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCAC  
 TTTGGAAGAGGTACAGGCAGAAAATCCCAAAGTGCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAG  
 CTTTACAGAGGGTCGCCATCCTCGTTCCCCACCGAACAGAGAGAAACACCTGATGTACCTGCTGGAA  
 CATCTGCATCCCTTCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGG  
 TAAAAAGTTTAATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG  
 ACTGCTTTATATCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAG  
 CATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTTTG  
 GGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAACTACTGGGGAT  
 GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTG  
 CCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG  
 GATGAAGCTCTTACACCAAGTGTACGAGTCTGGAGAACAGATGGGTGAGTAGTTGTTCTTATAAAT  
 TAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGGATTCTGGTTTGGTGCATGACCC  
 TGGATCTTTTGGTGATGTTTGAAGAAGTGAATCTTTGTTTGCAATAATTTGGCCTAGAGACTTCAA  
 ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTTGTTGAGCTGAATTTTTCCTTTTGTATTTCT  
 TAGCAGAGCTCCTGGTGATGTAGAGTATAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGAT  
 CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA  
 AATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTATGGGAT  
 AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA  
 AGGTACGAAGATACAATACTGTTATTCAATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGCAGGT  
 GAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTGGGAATGAAGA  
 GGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC  
 CTTCAAGGGAGGACCTGCCAGGTATGCCTTCCAGTGATGCCACCAGAGAATACATTCTCTATTAGT  
 TTTTAAAGAGTTTTTGTAAATGATTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT  
 ATTAATAATAATAATATGTCTATCAAAATACCTCTGTAGTAAAATGTGAAAAAGCAAAA

**FIGURE 42**

MGFNLTFFHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLGKGKTLTN  
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH  
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKGKFNRAKLLNVGYLEALKEENWDCFI FHDV  
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNGFSNNYWGWWGED  
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV  
SVEHNPLYINITVDFWFGA

**Important features:**

**Signal peptide:**

amino acids 1-27

**N-glycosylation sites.**

amino acids 4-8, 220-224, 335-339

**Xylose isomerase proteins.**

amino acids 191-202

# **FIGURE 43**

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGG  
 CCGCTTGCCCTCCTGCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCCACAACAG  
 ACGGGACAACCTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT  
 GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCTGCGGCTGCTGTC  
 ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCGTCCTCCCTCCC  
 TTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGAATAAAATGGCTGGTTCTTTTGT  
 TCCAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 44**

MALSSQIWAACLLLLLLLLSLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI  
CIFCCGCCHRSKCGMCCKT

**Important features:**

**Signal peptide:**

amino acids 1-24

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 58-59

**N-myristoylation site.**

amino acids 44-50

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 1-12

**FIGURE 45**

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC  
TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT  
TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG  
GACCTTCAACACAACCCCTCTTGTACCATAACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA  
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAAGTGAAG  
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA  
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA  
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT  
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC  
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTGAGCAGAAACT  
TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCCTCCATG  
GTCCCTCCTGTGTCTCCTGTTGGTGCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG  
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGACATTTGTCGGG  
AAACTCCTAACATATGCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT  
AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT  
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA  
TCTTAGACAGCAGTGCCTCCCTAAGTCTCTGCTCA

## **FIGURE 46**

MAGSPTCLTLIYILWQLTGAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQP  
 EGGTIIIVTQNRNRERVDFFDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK  
 PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC  
 VARNPVSRNFSSPILARKLCEGAADDPDSSMVLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIE  
 EKKRVDICRETPNICPHSGENTYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLTMPDT  
 PRLFAYENVI

**Important features:**

**Signal peptide:**

amino acids 1-22

**Transmembrane domain:**

amino acids 224-250

**Leucine zipper pattern.**

amino acids 229-251

**N-glycosylation sites.**

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,  
 291-295

**FIGURE 47**

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG  
ATTGAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT  
TAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTCCCAGGAATT  
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG  
CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTCAGTGTGATCACAGTCATTGGTGCTC  
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTCATGTGTAATTCTCCAAGC  
AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA  
CTTGCAAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTCAATAAACCCACCAGTAACG  
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAAACAAACATAGG  
CTTATCCACTTCTCAGTATTTTTAGGTCTATTGCTTGTGGAATTCTGGAGGTCCTGTTTGGGCT  
CAGTCAGATAGTCATCGGTTTCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG  
TGTAGTTTAAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAA

## **FIGURE 48**

MTCCEGWTSNCGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA  
 TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCLISIQALLKGPLMCNSPSNSNANCEFSL  
 KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSSEENKHRLIHFSVFLGL  
 LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

**Important features:**

**Transmembrane domains:**

amino acids 10-31 (type II), 50-72, 87-110, 191-213

**N-glycosylation sites.**

amino acids 80-84, 132-136, 148-152, 163-167

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 223-227

**N-myristoylation sites.**

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 207-218

**TNFR/NGFR family cysteine-rich region protein.**

amino acids 4-12



**FIGURE 49**

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA  
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG  
ATCCGTGGGCTGCAGACCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTCGAACTGTGA  
CATGGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGAAGCCAATGACC  
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC  
TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG  
CAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA  
CTACTTGCTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCAGCACC  
TCCTCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCCT  
TTCTGATCAGGAGGCTTCTTTATGAATTAACTCGCCCCACCACCCCTCA

**FIGURE 50**

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAAGIAAVLSGKCKYKS  
SQKQHSPVPEKAIPLITPGSATTC

**Important features:****Signal peptide:**

amino acids 1-16

**Transmembrane domain:**

amino acids 36-59

**N-myristoylation sites.**

amino acids 41-47, 45-51, 84-90

**Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.**

amino acids 54-67

**FIGURE 51**

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG  
 AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG  
 AGGAGACCTTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG  
 GGGCCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGGAGGCTGGCCCCCTGCAGAG  
 CGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGGACATGGCCTGGGAGACGCCCTGA  
 GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGGCAGCTGGCTCTAAAGTCAGTGAG  
 GCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC  
 AGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATGCTCTGGGAAACACTGGGCACGAGA  
 TTGGCAGACAGGCAGAAGATGTCTTCGACACGGAGCAGATGCTGTCCGCGGCTCCTGGCAGGGG  
 GTGCCTGGCCACAGTGGTGCTTGGGAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG  
 CCTTGGAGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCG  
 GAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGA  
 GGGCCACCAAACCTTTGGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTTCAGTGAG  
 AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA  
 ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC  
 AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG  
 TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT  
 CCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGA  
 CATAAACCCGGGTGTGAAAAGCCAGGGAATGAAGCCCGCGGAGCGGGGAATCTGGGATTCAGGG  
 CTTTCAAGGACAGGGAGTTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATCGCCTCCTTG  
 GAGGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTT  
 GGTGGAGTCAATACTGTGAACCTCTGAGACGTCTCCTGGGATGTTTAACTTTGACACTTTCTGGAA  
 GAATTTTAAATCCAAGCTGGGTTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGAAGCTCTC  
 GCATCCCGTGAACCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCACACTCCCTCCTTAA  
 AACACCACCTCTCATCTAATCTCAGCCCTTGCCCTTGAAATAAACCTTAGCTGCCCCACAAA  
 AA  
 AA



**FIGURE 53**

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG  
 CTGGGCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT  
 ACTCGCCCGCATCCTGGCTTGGACCTATGCCCTTCTATAACAACCTGCCGCCGGCTCCAGTGTTTCC  
 CACAGCCCCAAAACGGAACCTGGTTTTGGGGTCACCTGGGCCTGATCACTCTACAGAGGAGGGC  
 TTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGTATGGCTGGGTCCCAT  
 CATCCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA  
 TTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG  
 AGTGGCGGTGACAAGTGGAGCGGCCACCGTCGGATGCTGACGCGCGCCTTCCATTTCAACATCCT  
 GAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG  
 CCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA  
 CAGAAATGCATCTTCAGCTTTGACAGCCATTGTGAGGAGAGGGCCAGTGAATATATTGCCACCAT  
 CTTGGAGCTCAGTGCCCTTGTAAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT  
 ATTACCTCTCCCATGACGGGCGGCGCTTCCACAGGGCTGCCGCTGGTGCATGACTTCACAGAC  
 GCTGTATCCGGGAGCGGCGTCGCACCCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAA  
 AGCCAAGTCCAAGACTTTGGATTTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG  
 CATTGTGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCAG  
 GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG  
 ACAGGAGGTGCAAGAGCTTCTGAAGGACCGGATCCTAAAGAGATTGAATGGGACGACCTGGCCC  
 AGCTGCCCTTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATC  
 TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG  
 CCTCATCGATATTATAGGGGTCCATCACAACCCAACCTGTGTGGCCGGATCCTGAGGTCTACGACC  
 CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCCTTTCTCCGCA  
 GGGCCCAGGAACATGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCTGGCGTTGAT  
 GCTGCTGCACTTCCGTTTCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGC  
 GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCCTGAATGTAGGCTTGCAGACTTTCTGAC  
 CCATCCACCTGTTTTTTTGCAGATTGTTCATGAATAAAACGGTGCTGTCAA

**FIGURE 54**

MSLLSLPWLGLRPVAMSPWLLLLLVGSWLLARILAWTYAFYNNCRRRLQCFPQPPKRNWFWGHLG  
LITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP  
WLGEILLSGGDKWSRHRRLTPAFHFNIKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI  
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC  
RLVHDFTDVIRERRRTLPTQGIDDFKDKAKSKTLDLFDVLLSKDEDGKALSDEDIRAEADTF  
MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR  
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP  
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN  
VGLQ

**Important features:****Transmembrane domains:**

amino acids 13-32 (type II), 77-102

**Cytochrome P450 cysteine heme-iron ligand signature.**

amino acids 461-471

**N-glycosylation sites.**

amino acids 112-116, 168-172

# **FIGURE 55**

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGA  
 GCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTTCTGCCT  
 TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCAGTCTTTGGCATTGACGTGG  
 TACAGCCTTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTTTTGCCGTGTGTCTTGC  
ATAATTTCATGGCCAGTTTATGAAGCTTTGGAAGGCACTATGGACAGAAGCTGGTGACAGTTT  
 GTAACATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGACAGCAATGTGTTGCTT  
 GTGATTCGAACATTTGAGGGTTACTTTTGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA  
 GCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAATCTTCCTCATGTACCTGTTTCCTC  
 TCTGGATGTTGTCCCACTGAATTCCCATGAATACAAACCTATTCAGCAACAGCAAAAAAAAAAAAA  
 AA

## **FIGURE 56**

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGGLALIFCILQSLALTWYSLSFIPFAR  
DAVKKCFVCLA

**Important features:**

**Signal peptide:**

amino acids 1-33

**Type II fibronectin collagen-binding domain protein.**

amino acids 30-72



**FIGURE 57**

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCAGTGGAGCACCCAGCAGGCCGCCAACATGCTCTGTCTGTGCCTG  
 TACGTGCCGGTTCATCGGGGAAGCCAGACCGAGTTCCAGTACTTTGAGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCC  
 ATTTTCAAGCTCAGTGTCTTCATCCCTCCAGGAATTCTCCACCTACCGCCAGTGGAAAGCAGAAAAATTGTACAAGCT  
 GGAGATAAGGACCTTGATGGGCAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGAGG  
 CTGGTGTTTAAGATTTTGGACAAAAAAGATGATGGACGCATTGACGCGCAGGAGATCATGCAGTCCCTGCGGGACTTG  
 GGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTCTCAAGAGCATGGATAAAAAACGGCAGATGACCATCGACTGG  
 AACGAGTGGAGAGACTACCACCTCCTCCACCCCGTGGAAAAACATCCCCGAGATCATCCTCTACTGGAAGCATTCACG  
 ATCTTTGATGTGGGTGAGAATCTAACGGTCCCGGATGAGTTACAGTGGAGGAGAGGAGACGCGGGATGTGGTGGAGA  
 CACCTGGTGGCAGGAGGTGGGGCAGGGGCCGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG  
 CAGGTCCATGCCTCCGCGAGCAACAACATGGGCATCGTTGGTGGCTTCACTCAGATGATTGAGAAGGAGGGGCCAGG  
 TCACTCTGGCGGGCAATGGCATCAACGTCTCAAAATTGCCCCGAATCAGCCATCAAATTCATGGCCTATGAGCAG  
 ATCAAGCGCCTTGTGGTAGTGACCAAGGAGACTCTGAGGATTACAGAGAGGCTTGTGGCAGGGTCCCTGGCAGGGGCC  
 ATCGCCAGAGCAGCATCTACCCAATGGAGGTCTGAAGACCCGGATGGCGCTGCGGAAGACAGGCCAGTACTCAGGA  
 ATGCTGGACTGCGCCAGGAGGATCCTGGCCAGAGAGGGGGTGGCGCCTTCTACAAAGGCTATGTCCCCAACATGCTG  
 GGCATCATCCCCATGCCGGCATCGACCTTGCACTACGAGACGCTCAAGAAATGCCTGGCTGCAGCACTATGCAGTG  
 AACAGCGCGGACCCCGGCTGTTGTGCTCCTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGGCCAGCTAC  
 CCCCTGGCCCTAGTCAGGACCCGGATCGAGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCTC  
 TTCAAACATATCCTGCGGACCGAGGGGGCCTTCGGGCTGTACAGGGGGCTGGCCCCCACTTCATGAAGGTCAATCCCA  
 GCTGTGAGCATCAGCTACGTGGTCTACGAGAACCTGAAGATCACCTGGGCGTGCAGTCGCGGTGACGGGGGAGGGC  
 CGCCCGGCGAGTGGACTCGCTGATCCTGGGCGCAGCCTGGGGTGTGCAGCCATCTCATTCTGTGAATGTGCCAACCT  
 AAGCTGTCTCGAGCCAAGCTGTGAAAACCTAGACGCACCCGAGGGAGGCTGGGGAGAGCTGGCAGGCCAGGGCTT  
 GTCCTGCTGACCCAGCAGACCTCCTGTTGGTTCAGCGAAGACCACAGGCATTCCTTAGGGTCCAGGGTCAAGCAGG  
 CTCGCGGCTCACATGTGTAAGGACAGGACATTTCTGCAGTGCCTGCCAATAGTGAGCTTGGAGCCTGGAGGCCGGCT  
 TAGTTCCTCCATTTACCCCTTGACGCCAGCTGTTGGCCACGGCCCCCTGCCCTCTGGTCTGCCGTGCATCTCCCTGTGC  
 CCTCTTGCTGCCTGCCTGTCTGCTGAGGTAAAGGTGGGAGGAGGGCTACAGCCACATCCCACCCCTCGTCCAATCCC  
 ATAATCCATGATGAAAGGTGAGGTACGTGGCCTCCAGGCCCTGACTTCCCACTACAGCATTGACGCCAATTTGGC  
 TGTGAAGGAAGAGGAAAGGATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGGCATG  
 CTTGGGAGTGCAGGGGGCTCGGGCTGCCTGGCCTGGCTGCACAGAGGCAAGTGTGGGGCTCATGGTGTCTGAGCT  
 GGCTGGACCTGTGAGGATGGGCCCCACCTCAGAACCAAACTCACTGTCCCCACTGTGGCATGAGGGCAGTGGAGCA  
 CCATGTTTGAAGGCGAAGGGCAGAGCCTTTGTGTCTCTGGGAGGGAAGGAAAGGTGTTGAGGCCTTAATTATGG  
 ACTGTTGGGAAAAGGGTTTGTCCAGAAAGGACAAGCCGGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG  
 GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCTGTCCAACCCAGCAGGGGGCGCAGC  
 GGGACAGCCCCACATTCACCTGTGTCACTGCTTGGAACTTATTATTTGTATTTATTTGAACAGAGTTATGTCCT  
 AACATATTTTATAGATTTGTTAATTAATAGCTTGTCAATTTCAAGTTCATTTTATTTATCATATTTATGTTATGTT  
 GATTGTACCTTCCCAAGCCCGCCAGTGGGATGGGAGGAGGAGGAGAAGGGGGCCTTGGGCCGCTGCAGTCACATCT  
 GTCCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAGCGGCCAGAGGAGCAGCCCTGGCTCCTTTCTTTTGGCAG  
 GTTGGGGAAGGGCTTGGCCCCAGCCTTAGGATTTAGGGTTTGAAGTGGGGGGCTGGAGAGAGAGGAGGAACCTCAAT  
 AACCTTGAAGGTGGAATCCAGTTATTTCTGCGCTGCGAGGGTTTCTTTATTTCACTCTTTTCTGAATGTCAAGGCAG  
 TGAGGTGCTCTCACTGTGAATTTGTGGTGGGCGGGGCTGGAGGAGAGGGTGGGGGGCTGGCTCCGCTCCCTCCAGC  
 CTTCTGCTGCCCTTGTAAATGCGGGCCAATGGCGACCTCACGGTTGCATTCATTCCACCAGAATGACCTGA  
 TGAGGAAATCTCAATAGGATGCAAGATCAATGCAAAATTTGTTATATATGAACATATAACTGGAGTCGTCAAAAAG  
 CAAATTAAGAAAGAAATGGACGTTAGAAGTTGTCAATTAAGCAGCCTTCTAATAAAGTTGTTTCAAAGCTGAAAAA  
 AA

## FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSI FKL SVFIPSQEFSTYRQWKQKIVQAGDKDL DG  
 QLD FEEFVHYLQDHEKKLRLVFKILD KKN DGRIDAQEIMQSLRDLGVKISEQQA EKILKSMDKNG  
 TMTIDWNEWRDYHLLHPVENIPEI ILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRHLVAGGG  
 AGAVSRTCTAPLDR LKVL MQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESA IK  
 FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAI AQSSIYPMEVLKTRMALRKTGQYSGMLDCARR  
 ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKN AWLQHYAVNSADPGVFVLLACGTMSSTC  
 GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGA FGLYRGLAPNFMKVIPAVSISYV  
 VYENLKITLGVQSR

### Important features:

#### Signal peptide:

amino acids 1-16

#### Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

#### Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

#### N-glycosylation sites.

amino acids 129-133, 169-173

#### Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

## FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGC  
 TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG  
 CACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTACTGTCGCCTCAGCT  
 GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACCTTTTGAACCTGACATCAAACCTTTCTGATAT  
 CGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG  
 AGCTGTGCGAGCAGGATGAAATGTTTCAGAGGCCGGACAGCAGTGTTCGCTGATCAAGTGATAGTT  
 GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT  
 CATCACTTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG  
 AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCCC  
 CAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACCTCTCGGAAGTCTCCAATAC  
 CAGCTTTGAGCTGAACCTGAGAATGTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGA  
 TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG  
 ACAGAATCGGAGATCAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGT  
 CTCTTCTTTCTTTGCCATCAGCTGGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAAT  
 GTGCCTTGCCACAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTTAC  
 CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG  
 AGCAAACAAGAGCAAGAAACAAAAGAAGCCAAAAGCAGAAGGCTCCAATATGAACAAGATAAAT  
 CTATCTTCAAAGACATATTAGAAGTTGGGAAATAATTTCATGTGAACTAGACAAGTGTGTTAAGA  
 GTGATAAGTAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT  
 CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTTAGTTATATGTGC  
 TGTAATGTTGCTCTGAGGAAGCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCAC  
 AAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG  
 GCTGCATTTTAGTAATGGGTCAAATGATTCACCTTTTTATGATGCTTCCAAAGGTGCCTTGGCTTC  
 TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGT  
 CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTTAAACAAAAAAAAAAAAAAAAAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 60**

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS  
 DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASRLKNVQLTDAGTYKC  
 YIITSKGKGNANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS  
 NTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTSEIKRRSHLQLLNSKASL  
 CVSSFFAISWALLPLSPYMLK

**Important features:**

**Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 258-281

**N-glycosylation sites.**

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,  
 220-224

**N-myristoylation sites.**

amino acids 52-58, 126-132, 188-194

## FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACAAG  
CACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT  
GGGCTACCCCCTGGTGGTGGTTATGGGGGTCTGCCCCCTGGAGGGCCTTATGGACCACCAGCTGG  
TGGAGGGCCCTATGGACACCCAATCCTGGGATGTTCCCCTCTGGAACCTCAGGAGGACCATATG  
GCGGTGCAGCTCCCGGGGGCCCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGCCCAGCAGCCT  
GGGCTTTATGGACAGGGTGGCGCCCCCTCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC  
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA  
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA  
GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGGAAGAACCTCTT  
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC  
AAATGGGCTACAACCTGAGCCCCAGTTCACCCAGCTTCTGGTCTCCCGCTACTGCCCAGCTCT  
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA  
GGCCTTCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA  
CCATGACAGCTTCTCGGATGCTATGACCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT  
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTTCTGTCCCTCTAGAAGAAC  
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCTGCATCATAGCCACCA  
AATAGTGAGGACCGGGGCTGAGGCCACACAGATAGGGGCTGATGGAGGAGAGGATAGAAGTTGA  
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTTGTAATGG  
AGTTAGTGTCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGAGTGTTTCATCGGCCTGTTACC  
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTTCTCCAAAGTGGAAT  
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGGATTCTGCCACACCCATAAAT  
CCTTGTGTGTTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT  
CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCTGCCATGCTCTGCTCGGCT  
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTTAATTTGCATTTTTTTTC  
ATTTGGGGCCAAAAGTCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

**FIGURE 62**

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG  
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVSDSH  
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDR  
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK  
DTAVQGNIRLSFEDFVTMTASRML

**Important features of the protein:**

**Signal peptide:**

amino acids 1-19

**N-glycosylation site.**

amino acids 147-150

**Casein kinase II phosphorylation sites.**

amino acids 135-138, 150-153, 202-205, 271-274

**N-myristoylation sites.**

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-  
66, 70-75, 78-83, 83-88, 87-92, 110-115

## FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC  
 TCCCAGGCCTCTTTGCCCCGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCAAAACTTCGGGACC  
 AACTTGCCCTCAGCTCGGACAACCTTCCTCCACTGGCCCCCTCTAACTCTGAACATCCGCAGCCCCG  
 TCTGGACCCCTAGGTCTAATGACTTGGCAAGGGTTCTCTGAAGCTCAGCGTGCCTCCATCAGATG  
 GCTTCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG  
 GATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCCTGGGGGA  
 AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGCTGCGGCCCTCGCTCCGGGCAGTGGCCCTT  
 TGCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC  
 TCGGAGTCCAGACGACTGCCCCGTTCTAATTCACTGGGAGCCGGGGGAAAAATCCTTTCCCAACG  
 CCGTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGGTACCCTGAATCCCAGTG  
 TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCCATGCCACACCCTGAGGGA  
 ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG  
 CAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTAATCGGTATCCAGGAG  
 GCAGCTGGGGGAATATTCATCTATACCCAGGTATCAATAACCCATTTCCTCCTGGAGTTCTCCGC  
 CCTCCTGGCTCTTCTTGGAACATCCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG  
 GGGCTAGAGCACGATAGAGGGAAACCAACATTGGGAGTTAGAGTCCTGCTCCCGCCCCTTGCTG  
 TGTGGGCTCAATCCAGGCCCTGTAAACATGTTTCCAGCACTATCCCCACTTTTCAGTGCCTCCCC  
 TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
 AA

# THEORETICAL BACKGROUND

Important features of the protein:

amino acids 1-26

amino acids 56-59, 155-158

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,  
259-264, 269-274, 270-275, 280-285, 281-286, 305-310



## FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC  
TGGGCTGCCCCCTTGCTCCTCTTGACCCCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA  
CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA  
TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCTCCACCATGC  
AAGATCTCAACACCATGTTGTCTGCAACACATGACAGCCATTGAAGCCTGTGTCCTTCTTGGCCC  
GGGCTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAGCAGGCCCCCACCTC  
CTGAGTGGCAATAAATAAAATTCGGTATGCTG

## **FIGURE 66**

MSGSLPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSTSVTL  
HHARSQHHVVCNT

**Important features:**

**Signal peptide:**

amino acids 1-19

**N-glycosylation site.**

amino acids 37-41

**N-myristoylation sites.**

amino acids 15-21, 19-25, 60-66

## FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGGGC  
 CAGGTGCCCCGTCGCAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC  
 TTCTTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTGCTTCTGGCG  
 CTGGGCCTGCCGTTCTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC  
 AAATGAGAATAGCACTGTTTTGCCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG  
 AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG  
 GCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCCAGTAGCGA  
 GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT  
 GCCTGCCCATCTAGTCCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA  
 GGCAGTGGCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAA  
 AGACTCTGCCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACTTTATATATTTATATAAAATTAG  
 TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

[illegible]

Important features:

amino acids 1-19

amino acids 56-80

amino acids 36-40

amino acids 86-90

amino acids 86-94

amino acids 7-13, 26-32

**FIGURE 69**

GCCAGGAATAACTAGAGAGGAACAATGGGGTTATTAGAGGTTTTGTTTTCTTCTAGTTCTGTGCCTGCTGCACCAG  
TCAAATACTTCTTTTCAATTAAGCTGAATAATAATGGCTTTGAAGATATTGTCATTGTTATAGATCTAGTGTGCCAGAA  
GATGAAAAAATAATTGAACAAATAGAGGATATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGA  
TTTTTTTTTCAAAAATGTATCTATATTAATTCCTGAGAATTGGAAGGAAAAATCCTCAGTACAAAAGGCCAAAACATGAA  
AACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCATACCAAGCAGTTTCA  
GAATGTGGAGAGAAAGGCCAATACATTCACTTCACCCCGACCTTCTACTTGGAAAAAACAAAATGAATATGGACCA  
CCAGGCCAACTGTTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGTGATGAGTACAATGAAGATCAGCCTTTC  
TACCGTGCTAAGTCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAAGTGT  
CAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAACGTATGGAAAAGATTGTCAATTCTTT  
CCTGATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATGCAAAGTATTGATTCTGTTGTTGAATTTTGTAAACGAA  
AAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAAT  
TCTGAGGATTTTAAAAACACCATACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATCAGTCAA  
AGAATTTGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGGTAAAGGACCGCTAAATCGAATGAATCAAGCAGCA  
AAACATTTCTGCTGCAGACTGTTGAAAATGGATCCTGGGTGGGGATGGTTCACTTTGATAGTACTGCCACTATTGTA  
AATAAGCTAATCCAATAAAAAAGCAGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCTCTGGGAGGA  
ACTTCCATCTGCTCTGGAATTAATATGCATTTCAAGGTGATTGGAGAGCTACATTTCCCACTCGATGGATCCGAAGTA  
CTGCTGCTGACTGATGGGAGGATAAAGTCTGCAAGTTCTTGATGATGAGTGAACAAAGTGGGGCCATTGTTTATGTT  
TTTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTTTATGTT  
TCAGATGAAGCTCAGAACATGGCCTCATTGATGCTTTTGGGGCTTTACATCAGGAAATACTGATCTCTCCAGAG  
TCCCTTCAGCTCGAAAGTAAGGGATTAACTGAATAGTAATGCCTGGATGAACGACACTGTCAATATGATAGTACA  
GTGGGAAGGACACGTTCTTTCTCATCATGGAACAGTCTGCCTCCAGTATTCTCTCTGGGATCCAGTGGGAACA  
ATAATGGAAAAATTTACAGTGGATGCAACTTCCAAAATGGCCTATCTCAGTATTCCAGGAAGTCAAAGGTGGGCACT  
TGGGCATACAATCTTCAAGCCAAAGCGAACCCAGAAACATTAATATACAGTAAGTCTCGAGCAGCAATTTCTTCT  
GTGCTCCAATCACAGTGAATGCTAAATGAATAAGGACGTAAACAGTTTCCCGAGCCCAATGATTGTTTACGCAGAA  
ATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTTCAATGAATCACAGAATGGACATACAGAAGTT  
TTGGAACTTTTGATAATGGTGAGGCGCTGATTCTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATAT  
ACAGAAAATGGCAGATATAGCTTAAAGTTTCGGGCTCATGGAGGAGCAAACTGCCAGGCTAAATACGGCCTCCA  
CTGAATAGAGCCGCTACATACCAGGCTGGGTAGTGAACGGGGAAATGAAGCAACCCGCCAAGACCTGAAATTTGAT  
GAGGATCTCAGACCACCTTGGAGGATTTAGCCGAACAGCATCCGGAGGTGCTTTTGTGGTATCACAAGTCCCAAGC  
CTTCCCTTGGCTGACCAATACCCACCAAGTCAATCACAGACCTTGATGCCAGTTTATGAGGATAAGATTATTTCTT  
ACATGGACAGCACCAGGAGATAATTTTATGTTGGAAGTTCAACGTTATATCATAAGAATAAGTGCAAGTATTCTT  
GATCTAAGAGACAGTTTGTATGATGCTCTTCAAGTAAATACTACTGATCTGTACCAAAAGGAGGCCAATCCAAGGAA  
AGCTTTGCATTTAAACAGAAAAATATCTCAGAAGAAAAATGCAACCCACATATTTATGGCATTAAAAAGTATAGATAAA  
AGCAATTTGACATCAAAAGTATCCAACATTGCACAAGTAACTTTGTTTATCCCTCAAGCAATCCTGATGACATTGAT  
CCTACACCTACTCCTACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATTTTCTACGCTGGTATTG  
TCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTAAGTACCACCATTTGAACCTTAACGAAGAAAAAATCTTC  
AAGTAGACCTAGAAGAGAGTTTTAAAAACAAAAAATGTAAGTAAAGGATATTCTGAATCTTAAATTCATCCCAT  
GTGTGATCATAAACTCATAAAATAATTTAAGATGTGGAAGGAGTACTTTGATTAAATAAAAAACACTCATGGATA  
TGTA AAAACTGTCAAGATTAATAATTAATAGTTTCATTTATTTGTTATTTTATTGTAAGAAATAGTGATGAACAAAG  
ATCCTTTTTCATACTGATACCTGGTTGATATTTTGTATGCAACAGTTTCTGAAATGATATTTCAAATTCATCAA  
GAAATTAATCATCTATCTGAGTAGTCAAAATACAAGTAAAGGAGAGCAATAAACCAACATTTGGAAAAA  
AA

**FIGURE 70**

MGLFRGFVFLVLCLLHQSNTSFIKLNNNGFEDIVIVIDPSVPEDEKIEQIEDMVTTASTYLFE  
 ATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY  
 IHFTPDLLLGGKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN  
 RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVQTEKASIMFMQSIDSVVEFCNEKTHNQEAP  
 SLQNIKCNRSTWEVISNSEDfKNTIPMVTPPPPVFSLLKISQRIVCLVLDKSGSMGGKDRLNR  
 MNQAAKHFLLOTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK  
 YAFQVIGELHSQLDGSEVLLLTGDNATSSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG  
 SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFL  
 ITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR  
 AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA  
 GADSFKNDGVYSRYFTAYTENGRYSLKVRHGGANTARLKLRLPLNRAAYIPGWVVNGEIEANPP  
 RPEIDEDTQTTLDEFRTASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIIILTWTAPGDN  
 FDVGKVQRYIIRISASILDLRDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI  
 KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTLVLSVIGSVVI  
 VNFILSTTI

**Signal peptide:**

amino acids 1-21

**Putative transmembrane domains:**

amino acids 284-300, 617-633

**Leucine zipper pattern.**

amino acids 469-491, 476-498

**N-glycosylation site.**
 amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,  
 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

**FIGURE 71**

CTCCTTAGGTGGAAACCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCTCCCGGGCAGGGGTGA  
CAACAGGTGTATCTTTTGTCTCGTGTGGCTGCCTTCTATTTCAAGGAAAGACGCCAAGGTAATTTTGACCCA  
GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCAAGTTATGCCAGGATTTACTAGAGAGTGTCA  
ACTCAACCAGCAAGCGGCTCCTTCGGCTTAACCTGTGGTTGGAGGAGAGAACCCTTGTGGGGCTGCGTTCTCTTAGCA  
GTGCTCAGAAAGTGACTTGCCTGAGGGTGGACCAGAAGAAAGGAAAGGTCCCCCTTGTGCTGTGGCTGCACATCAGGAA  
GGCTGTGATGGGAATGAAGGTGAAACTTGGAGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCTTTAAA  
AGTAGAGAAGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC  
CGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTGGGGGGCCCCGGCTTTGAGGGATGCC  
ACCGGTTCTGGACGCATGGCTGATTCTGAATGATGATGGTTTCCGCGGGGGCTGCTTGCCTGGATTTCCTGGGTGGTG  
GTTTTGTGGTGTCTCTGTCTGTGCTATCTGTCTGTACATGTTGGCTGCACCCCAAAAGGTGACGAGGAGCAG  
CTGGCACTGCCAGGCCAACAGCCCCACGGGGAAGGAGGGGTACCAGGCCCTCCTTCAGGAGTGGGAGGAGCAGCAC  
CGCAACTACGTGAGCAGCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGCAGCTCAGG  
AATGGGCAGTACCAAGCCAGCGATGCTGCTGGCTGGGTCTGGACAGGAGCCCCCAGAGAAAAACCCAGGCCGACCTC  
CTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCTCAAGCTGGCCACAGAGTATGCAGCAGTG  
CCTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGACTGGCCTTACCCGCCACCCGAGGAGAAGCCTGTG  
AGGAAGGACAAGCGGGATGAGTTGGTGGAAAGCCATTGAATCAGCCTTGGAGACCCTGAACAATCTGCAGAGAACAGC  
CCCAATCACCCTCCTTACAGCGCTCTGATTTCATAGAAGGGATCTACCGAACAGAAAGGACAAAGGCACATTGTAT  
GAGCTCACCTTCAAAGGGGACCACAAACACGAATTCAAACGGCTCATCTTATTCGACCAATTCAGCCCCATCATGAAA  
GTGAAAAATGAAAAGCTCAACATGGCCCAACACGCTTATCAATGTTATCGTGCCTTAGCAAAAAGGGTGGACAAGTTC  
CGGCAGTTCATGCAGAAATTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGG  
AAAGAAGAAATAAATGAAGTCAAAGGAATACTTGAAAACACTTCCAAAGCTGCCAACTTCAGGAACCTTCACTTCATC  
CAGCTGAATGGAGAATTTTCTCGGGGAAAGGGACTTGATGTTGGAGCCCGCTTCTGGAAGGGGAAGCAAGCTCCTTCTC  
TTTTTCTGTGATGTGGACATCTACTTCACTCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAG  
GTATTTTATCCAGTTCCTTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGGAA  
CAGCAGCTGGTCATAAAGAAGGAACTGGATTTTGGAGAGACTTTGGATTGGGATGACGTGTCAATCTCGGTGAGAC  
TTCATCAATATAGGTGGGTTGATCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTC  
CACAGCAACCTCATAGTGGTACGGACGCTGTGGAGGACTTCCACCTCTGGCATGAGAAGCGGTGCATGGACGAG  
CTGACCCCGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCACGGCCAGCTGGGCATGCTG  
GTGTTTCAGGCACGAGATAGAGGCTCACCTTCGCAACAGAAACAGAAAGCAAGTAGCAAAAAACATGAAGTCCCGAGA  
GAAGGATTGTGGGAGACACTTTTCTTCTCTTTCGAATTAAGTGGCTGCAACAGAGAAAAGACTTCCATAAA  
GGACGACAAAAGAAATGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTCTCTCTGTTGGGCTTTTACAAACAGA  
AATCAAAATCTCCGCTTTCCTGCAAAAGTAACCCAGTTGCACCTGTGAAGTGTCTGACAAAGGCAGAAATGCTTGTG  
AGATTATAAGCCTAATGGTGTGGAGGTTTGTATGGTGTATTAATACACTGAGACCTGTTGTTTTGTGTGCTCATTTGA  
AATATTCATGATTTAAGAGCAGTTTGTAAAAAATTCATTAGCATGAAAGCAAGCATATTTCTCCTCATATGAATGA  
GCCTATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAAATATCAGAAGGCAGGAGAGGAGATAGGCTTATTATGATACT  
AGTGAGTACATTAAGTAAAAATAAATGGACCAGAAAAGAAAAGAACATAAATATCGTGTCAATTTTCCCAAGAT  
TAACCAAAATAATCTGCTTATCTTTTGGTTGTCTTTTAACTGTCTCCGTTTTTTCTTTTATTTAAAAATGCACT  
TTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTGCAAGCCTTACAAGAGACACAAGTTGGCCTAC  
ATTTTTATATTTTTTAAGAAGATACTTTGAGATGCATTATGAGAACTTTCAGTTCAAAGCATCAAATGTATGCCATAT  
CCAAGGACATGCCAAATGCTGATTCTGTGAGGCACTGAATGTGAGGCATTGAGACATAGGGAAGGAATGGTTTGTACT  
AATACAGACGTACAGATACTTCTCTGAAGAGTATTTTCAAGAGGAGCAACTGAACACTGGAGGAAAAAGAAATGAC  
ACTTCTGCTTTACAGAAAAGGAACTCATTGAGACTGGTGATATCGTGATGATCTAAAGTCAAGAAACACATTTT  
CTCCTCAGAAAGTAGGGACCGCTTCTTACCTGTTTAAATAAACCAAAGTATACCGTGTGAACCAACAATCTCTTTT  
AAAAAGGGTGTCTCTCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAAATATATATATATATATATATTTGT  
GAAAGATCAATCCATCTGCCAGAACTAGTGGGATGGAAGTTTTTGTACATGTTATCCACCCAGGCCAGGTGGAAG  
TAACTGAATATTTTTTAAATAAGCAGTCTTACTCAATCACCAGATGCTTCTGAAAATTGCAATTTTATTAACATTT  
CAAACTATTTTTTAAAAATAAATACAGTTAATACATAGAGTGGTTTCTTCAATCATGTGAAAATTTATAGCCAGCACCAG  
ATGCATGAGCTAATATCTCTTGAAGTCTGCTTCTGTTTGTGCTCACAGTAAACTCATTTGTTTAAAAAGCTTCAAGAAC  
ATTCAAGCTGTTGGTGTGTTAAAAATGCATTGTATTGATTGTACTGGTAGTTTATGAAATTTAATTAACACAGG  
CCATGAATGGAAGGTGGTATTGCACAGCTAATAAATATGATTGTGGATATGAA

**FIGURE 72**

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEW  
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK  
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETNNPA  
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNEKLNMAN  
TLINIVVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF  
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNLCRLNTQPGKKVFYFVLF  
SQYNPGIIYGHHDVAPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDDIKGWGGEDVHLYR  
KYLHNSLIVVRTFVRGLFHLWHEKRCMDELTPQYKMCMSKAMNEASHGQLGMLVFRHEIEAHL  
RKQKQKTSSKKT

**Important features:****Signal peptide:**

amino acids 1-27

**N-glycosylation sites.**

amino acids 315-319, 324-328

**N-myristoylation sites.**

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

**Amidation site.**

amino acids 377-381





## **FIGURE 74**

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHPASYPECGNNAQSPIDIQTDSVTDFDLPALQ  
 PHGYDQPGTEPLDLHNNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKSPGGSEHQINSEAT  
 FAELHIVHYDSDSYDSLSEAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP  
 ENLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV  
 QNYRALQPLNQRMVFASFQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLNRKS  
 VVFTSAQATTEA

**Important features of the protein:**

**Signal peptide:**

amino acids 1-15

**Transmembrane domain:**

amino acids 291-310

**N-glycosylation site.**

amino acids 213-216

**Eukaryotic-type carbonic anhydrases proteins**

amino acids 197-245, 104-140, 22-69

**FIGURE 75**

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCTGTGTC  
 TCTGGTGGTTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACA  
 ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTATCACAA  
 ATTGGCCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCC  
 AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAATATACTCCAATCTGA  
 AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC  
 ACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCTTCGTCCC  
 AGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT  
 CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTCCTTTTT  
 TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT  
 GATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCGCTGCTGAAAAAATCGTGATTA  
 ACTTTATCACCCCTCAATATCTCGGATGATTCTAAAAATTTCTCATCAGGATATGAGTTTACTGGGA  
 AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA  
 GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAG  
 AAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAATACCCCGGATAAA  
 ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTTGTGCGGGGCCTGAAGAGCAGGA  
 GCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGAGGCAGCGTTGGCAG  
 TCTTGGGCCCGCAAACGTTACAGTACTCATAACCCCTCAGCTCCAAGACTTAGACCCCTGGCG  
 CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGA  
 TCCCCAAACTGGCAGGCTGTGTATTCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCG  
 AGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT  
 CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGTTATA  
 TGTGCAGATGGAAAACTGATGCCAACACTTCCTTTTGCCTTTTGTTCCTGTGCAAACAAGTGAG  
 TCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTGAGTGT  
 CTGTGAGAATTACTTATTTCTTTCTCTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTA  
 GGTCTCTTAACAATGATGGTGGGCCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA  
 AGCAGTCAATAAATGTTTGGCAGACTGGGTGCAGAATTTATTTCAGGTGGGTGT

## **FIGURE 76**

MSYNGLHQRVFKELKLLTLCSSISQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY  
 SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSÉKQCARTLK  
 DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNÉFDKRFFVPAEK  
 IVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC  
 DSEENTEGTSLTQQESLSRTIPDKTVIEYDYVRTTDICAGPEEQELSLQEEVSTQGTLLSQA  
 ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEÉPSTTLVDWDPQTGRLCIPSLSSFDQDS  
 EGCEPSEGDGLGEEGLLSRLYEÉPAPDRPPGENETILMQFMEEWGLYVQMEN

### **Important features:**

#### **Signal peptide:**

amino acids 1-28

#### **Transmembrane domain:**

amino acids 140-163

#### **N-glycosylation sites.**

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

**FIGURE 77**

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC  
 CTGGGAAGATGGCCGGCCCGTGGACCTTCACCCCTCTCTGTGGTTTGCTGGCAGCCACCTTGATC  
 CAAGCCACCCTCAGTCCCACTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC  
 ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC  
 GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCTGAAGCACATC  
 ATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA  
 GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA  
 TCGTGGAGTTCCACATGACGACTGAGGCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC  
 CCCACCCGCTGGTCCCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAAGTGTGTA  
 TAAGCTCTCCTTCTGGTGAACGCCTTAGCTAAGCAGGTGATGAACCTCCTAGTGCCATCCCTGC  
 CCAATCTAGTGAAAACAGCTGTGTCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC  
 CTCTGCAGCTGGTGAAGGTGCCATTTCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTA  
 TCCTGCCATCAAGGTGACACCATTAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAA  
 AGGTGACCAAGTGTTCAATAACTCTGCAGCTTCCCTGACAATGCCACCCTGGACAACATCCCG  
 TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA  
 ATTCATGGTCTGTGGACTCTGTGCTTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC  
 TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC  
 ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTT  
 TCCCTCCAGTGAAGCCCTCCGCCCTTTGTTACCCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT  
 TTACACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG  
 ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA  
 CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG  
 CCTTGGGATTGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCC  
 TTGTGGAAACCCAGCTCTCTGTCTCCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAGGCTGG  
 GTCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACTTG  
 CCTGTGAAAAA

## **FIGURE 78**

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIEKLTQELKDHNATSILQQLPLLSAMREK  
PAGGIPVLGSLVNTVLKHI IWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE  
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLENL  
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDITQLYLGAQLLDLSQKVT  
KWFNNSAASLTMPITLDNIPFSLIVSQDVVKAAVAVALSPEEFMVLLDSVLPESAHLKSSIGLIN  
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPLFTLGIEASSEAQFYT  
KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSVLKALG  
FEAAESSLTkdalvltpaslwkpsspvsq

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-21

#### **N-glycosylation sites.**

amino acids 48-51, 264-267, 401-404

#### **Glycosaminoglycan attachment site.**

amino acids 412-415

#### **LBP / BPI / CETP family proteins.**

amino acids 407-457

# **FIGURE 79**

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC  
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT  
CTTGGCCTCCAAC TTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT  
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGTTGGCTTCT  
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC  
ACCCCTTCTGGGCCTGCCCCTGACATCCAGGCTGCCCAGGCCATGATGGTGACATCCAGTGCAAT  
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC  
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTCATCCTTGGAGGCCTCCTGGGATTC  
ATTCTGTGTGCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCAC TGGTGCTGACAG  
CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG  
CTGGAATCATCCTCTGCTTTTCCTGCTCATCCCAGAGAAATCGCTCCAATACTACGATGCCTAC  
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA  
GTTCAATTCCTACAGCCTGACAGGGTATGTGTGAAGAACCAGGGGCCAGAGCTGGGGGGTGGCTG  
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT  
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTTGCCATTGGATTGAGCAAAGGCAGAAATGGGG  
GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTCTGTTTTCC  
TCACCTTGCTGCTCCCTGCCCTAAGTCCCCAACCCCTCAACTTGAAACCCCATTCCTTAAGCCA  
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA  
TCCCACTGACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCAT  
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC  
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCACTCTTGTATGACTCCACAGTGTCCA  
GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG  
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

## **FIGURE 8o**

MASLGLQLVGYILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD  
IYSTLLGLPADIQAAQAMMTSSAIISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL  
LGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY  
DAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-24

#### **Transmembrane domains:**

amino acids 82-102, 117-140, 163-182

#### **N-glycosylation site.**

amino acids 190-193

#### **PMP-22 / EMP / MP20 family proteins.**

amino acids 46-59





## **FIGURE 82**

MVPEVRVLSSLLGLALLWFPLDSHARAREDMFCLFHGKRYS PGESWHPYLEPQGLMYCLRCTCSE  
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCEPHTPSGLRAPPKSCQHNGTMYQHGEI FSAHELF  
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR  
HPQDPCSSDAGRKRGPPTAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTY S  
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTTEYPCRHPEKVAGKCKKICPEDKADPGHSE  
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH  
SQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSPDGAEGHGQSRQSDQDITKT

### **Signal peptide:**

amino acids 1-25

## FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCAGCGAGAGCCTCTCC  
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTCTCTCTAATCCATCCGTCACCTCTCCTGTCA  
TCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTTGTTCTGAGTC  
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCTTGTTGGGGGAG  
GACGCAGCATTTCTCCTGTTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGG  
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTATGCAGATGCCACAGTATC  
AAGGCAGGACAAAACGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTACT  
GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCAGTCTTACTACCAGAAGGCCATCTGGGAGCT  
ACAGGTGTGAGCACTGGGCTCAGTTCTCTCATTTCCATCACGGGATATGTTGATAGAGACATCCAGCTAC  
TCTGTGAGTCTCGGGCTGGTTCCCCCGGCCACAGCGAAGTGGAAGGTCCACAAGGACAGGATTTGTCC  
ACAGACTCCAGGACAAACAGAGACATGCATGGCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAA  
CGCCGGGAGCATATCTGTTCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG  
GAGATACCTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGCCTA  
TTTTTTGGCATTGTTGGACTGAAGATTTTCTTCTCCAAATCCAGTGGAATACTCCAGGCGGAAGTGGACTG  
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACACGCAGTGAGGTGACTCTGGATCCAG  
AGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAAGTGTAAACCATAGAAAAGCTCCCCAGGAGGTG  
CCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGAAACATTA  
CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCGTGGGAGTGTGCCGGGATGATGTGGACAGGAGGA  
AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTTGTATTTT  
ACATTAAATCCCCGTTTATCAGCGTCTTCCCCAGGACCCACCTACAAAAATAGGGGTCTTCTGGACTA  
TGAGTGTGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCTGACATGTGGTTTG  
AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAAATGGAATCCCATAGTCATCTGC  
CCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGCAATCCAGAGACAAGCAACAG  
TGAGTCTCTCTCACAGGCAACACGCCCTTCTCCCCAGGGGTGAAATGTAGGATGAATCACATCCACAT  
TCTTCTTTAGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGCCAAGGTGGCTTCCA  
GATGAAGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCTGGCTGCCCTGAGCTGGGAGGGAAGAAGG  
CTGACATTACATTTAGTTTGTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAG  
AACCGTCAGGAATCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG  
ATCTTATTGATGACAGAGTGTATCCTAATGGTTTGTTCATTATATTACACTTTCAGTAAAAAAA

## **FIGURE 84**

MALMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSVVH  
 LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYQKAIWELQ  
 VSALGSVPLISITGYVDRDIQLLCQSSGWFPRTAKWKGPQGDLSTDSRTNRDMHGLFDVEISL  
 TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK  
 FQWKIQAELDWRRKHGQAEIRDARKHAVEVTLPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF  
 TRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVLRNLNGEHLFT  
 LNPRFISVFPRTPTTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPHYIEYPSYNEQNGT  
 PIVICPVTQESEKEASWQRASAIPESTSNESSSQATTPLPRGEM

**Signal peptide:**

amino acids 1-17

**Transmembrane domain:**

amino acids 239-255

**FIGURE 85**

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCCT  
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCTGA  
CGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT  
TACCCTGGCCCAGTAGTTCATGGCTACTGGTTCCGGAAGGGGCCAATACAGACCAGGATGCTCC  
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG  
GGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA  
TACTTCTTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA  
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCAGGCACCTGGAGTCCGGCTGCCCCC  
AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA  
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCTCATCCACACA  
GCCCCAGGACCATGGCACCAGCCTCACCTGTGAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA  
ACAAGACCGTCCATCTCAACGTGTCTACCCGCCTCAGAACTTGACCATGACTGTCTTCCAAGGA  
GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTCTCT  
GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA  
GAGGCCTGACCCTGTGCCCTCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCAC  
CTGAGGGATGCAGCTGAATTCACTGCAGAGCTCAGAACCCCTCTCGGCTCTCAGCAGGTCTACCT  
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGGAGCTGGAG  
CCACAGCCCTGGTCTTCTGTCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAAA  
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTGAGGGGTTT  
AGCCTCTCAGGGGCCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCAG  
CTTCTGCCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG  
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG  
ATGAGAAACTGCAGAGACTCACCTGATTGAGGGATCAGAGCCCTCCAGGCAAGGGAGAAGTCA  
GAGGCTGATTCTTGTAGAATTAAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG  
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCC  
TCCCTTTTATTTTTTAACTAAAAGACAGACAAATTCCTA

## FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIYPGPVVHGYWFREG  
 ANTQDAPVATNNPARAVWEETDRFHLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGS IKWNY  
 KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTTPMISWIGTSVSPDPSTTRS  
 SVLTLPQPQDHGTS LTCQVTFPGASVTNKT VHLNVSYPQNLMTVTFQGDGT VSTVLGNGSSL  
 SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP  
 LGSQQVYLVNLSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGD TGIE  
 DANA VRGSASQGPLEPWAEDSPPDQPPPASARSSVGE GELQYASLSFQMVKPWDSRGQEATDTE  
 YSEIKIHR

**Signal peptide:**

amino acids 1-15

**Transmembrane domain:**

amino acids 351-370

**FIGURE 87**

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC  
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC  
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT  
TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTTGA  
TGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG  
GGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG  
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC  
CAACTACAACACCTTTGGATCTGCAGAGGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT  
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCCATGCAGACTGG  
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT  
GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC  
CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCTAT  
GGCCAGCGGAATTCAGTGCGGGATTTGTTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA  
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG  
GATACTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT  
GGAACATCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG  
TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAA  
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA  
AAAAAA





**FIGURE 89**

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCAGCCTCAGAGAC  
CGCCGCCCTTGTCCCGAGGGCCATGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG  
CTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTG  
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTTCAGCTGGTGGCCGCGCTCT  
CTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCTCTCAGGAGTCTCCATGTTCAAC  
AGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT  
ATTCGAGCGTTGGGAGTGCACTACGTATTGGTACATTTTGTCTTCTGCAGTGCCCTTCCAGCTG  
TCACTGAAATGGCTTTATTCGTCACCGTCTTTGGGCTGAAAAGAAACCTTCTTGATTACCTTCA  
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCTGGAAGAAGGAAG  
GCATAGGCTTCGGTTTTCCCTCGGAAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG  
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGC'TTGTAAATAAAATATGTTTTGTAGTAACA  
TTAAGACTTATATACAGTTTtaggggacaattaaaaaaaaaaaa

**FIGURE 90**

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPFTTPEEYDKQDIQLVAALSVTLGLFA  
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCALPAVTEMALFV  
TVFGLKKKPF

**Transmembrane domain:**

amino acids 12-28 (type II), 51-66, 107-124

## FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC  
 TCGCTGCTGCTTCGTGTTCTTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG  
 GTCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT  
 CGGAAGCGGGGCCACATCTACCTAAGTCCCGCCCCATGGCCAATTCCACTCTCCTAGGGCTGCT  
 GGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCGAACCACAGCCCCC  
 CACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG  
 GCCCTGAACCTGCTCGTCCAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT  
 CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG  
 AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG  
 GAGTGGGAGAAGGTAGAACGGGGCCGCCGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTG  
 CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTTCAAAGTCGTCTGTG  
 TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACCTAC  
 CATAGTGATACCCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC  
 TGGAAGGACAGGCCTGCCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTTGGGCCTC  
 AGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG  
 AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG  
 AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG  
 TCCCCGAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC  
 CTTTGGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACA  
 GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGG  
 GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCCTTGTCGTGTGCTGAGCATGG  
 CATGAGGCTGAAGTGGAACCCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC  
 CAGGCCACCCCTTTCCAAAATTCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTTGCTGATG  
 GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCTCCACACTAAGGCCACAGCCCATC  
 CGCGTGCTGTGTGTCCTCTTCCACCCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCC  
 GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTGACACCGGGGTTCTCCCGGATCTGGATGGCGC  
 CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGCCGAGAGCATGTGCTGGATCTGTTT  
 TGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGGAAGTCTGTGAAACCGCTGATTGCTGACTTT  
 TGTGTGAAGAATCGTGTTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

## **FIGURE 92**

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL  
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF  
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRRTSLCTHDP  
AKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVC PDYNYHSDTPYYPSG

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-14

#### **N-glycosylation sites.**

amino acids 62-65, 127-130, 137-140, 143-146

#### **2-oxo acid dehydrogenases acyltransferase**

amino acids 61-71

**FIGURE 93**

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCCTGCGCTCGCC  
CTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGCTTTCTT  
CTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA  
AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTCTATATCCAAGAA  
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC  
AGGTGAGACAGCACCTCTATGCGACTGCTGGCCTATGTTTCTGGCTGGGCTTTGGAATCATGA  
GTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTTCAT  
GGAGATTCTCCTCAATTCTTCCTTTATTCAGCTTTCATGACGCTGGTCATTATCTTGCTGCATGT  
ATTCTGGGGCATTGTATTTTTTGATGGCTGTGAGAAGAAAAAGTGGGGCATCCTCCTTATCGTTC  
TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG  
TCAGCATTTATAATCCTGGTGCTCATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG  
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAACTTTCTTTTACAACCAGCGCTCCAGAT  
AACTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCT  
TTTTCTGAAAATCCCTTTTTCTGGTGGAATTGAGAAAGAAATAAACTATGCAGATA

## **FIGURE 94**

MTAAVFFGCAFIAFGPALALYVFTTIAIEPLRIIFLIAGAFFWLVSLLISLVWFMARVIIDNKDG  
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLAYVSGLGFGIMSGV  
FSFVNTLSDSLPGPTVGIHGDSPQFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKWGILLIVLLT  
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSCLKCLLCQDKNFLLYNQRSR

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-19

#### **Transmembrane domains:**

amino acids 32-51, 119-138, 152-169, 216-235

#### **Glycosaminoglycan attachment site.**

amino acids 120-123

#### **Sodium:neurotransmitter symporter family protein**

amino acids 31-65



# **FIGURE 96**

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM  
 LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPFI FVTQLGAQGTTLSSEE  
 LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG  
 TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

**Signal peptide:**  
 amino acids 1-16



## FIGURE 97

GCTCAAGTGCCTGCCTTGCCCCACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT  
CTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCTGACCATG  
GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCCAGGCTCTCCCCAAGGCCAGCCTGC  
AGAGCTGTCTGTGGAAGTTCAGAAAACATATGGTGGAAATTTCCCTTTATACCTGACCAAGTTGC  
CGCTGCCCCGTGAGGGGGTGAAGGCCAGATCGTGCTGTCAGGGGACTCAGGCAAGGCAACTGAG  
GGCCCATTTGCTATGGATCCAGATTCTGGCTTCTGCTGGTGACCAGGGCCCTGGACCGAGAGGA  
GCAGGCAGAGTACCAGCTACAGGTACCCCTGGAGATGCAGGATGGACATGTCTTGTGGGGTCCAC  
AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTTCTCTCAAGCCATCTAC  
AGAGCTCGGCTGAGCCGGGTACCAGGCTGGCATCCCCCTTCTCTTCTTGGGCTTCAGACCG  
GGATGAGCCAGGCACAGCCAACCTCGGATCTTCGATTCCACATCCTGAGCCAGGCTCCAGCCCAGC  
CTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC  
ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGA  
CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAAAGTCTCCATCATAGAGAGCAGCTGGGTGT  
CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTATACCCGACCCACATGGCCCAGGTA  
CACTGGAGTGGGGGTGATGTGCACTATCACCTGGAGAGCCATCCCCGGGACCTTTGAAGTGAA  
TGCAGAGGGAAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCAGGCTGAGTACCTGCTCC  
AGGTGCGGGCTCAGAATTCCTATGGCGAGGACTATGCGGCCCTCTGGAGCTGCACGTGCTGGTG  
ATGGATGAGAATGACAACGTGCCTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT  
CAGTCCACCAGGTACTGAAGTGACTAGACTGTGAGCAGAGGATGCAGATGCCCCGGCTCCCCCA  
ATTCCCACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTC  
CAGGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACAT  
CCTGCTTCTGGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCAGCTGTGAAG  
TCGAAGTCCAGTCACAGATATCAATGATCACGCCCTGAGTTTCATCACTTCCAGATTGGGCCT  
ATAAGCCTCCCTGAGGATGTGGAGCCCGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA  
CCTCGAGCCCGCTTCCGCTCATGGATTTTGCCATTGAGAGGGGAGACACAGAAGGGACTTTTG  
GCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG  
GCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGGCCAGGCC  
AGGCCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGGAGAGAGTGATGCCACCCCCAAGTTGG  
ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCCCAGCCGGCTCTTTCCTGCTGACCATC  
CAGCCCTCCGACCCCATCAGCCGAACCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCT  
CTGCATTGAGAAATTTCCGGGGAGGTGCACACCGCCAGTCCCTGCAGGGCGCCAGCCTGGGG  
ACACCTACACGGTGCTTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCCTGTGCCCTCCCAA  
TACCTCTGCACACCCCGCAAGACCATGGCTTGATCGTGAGTGGACCCAGCAAGGACCCCGATCT  
GGCCAGTGGGCACGGTCCCTACAGCTTACCCTTGGTCCCAACCCACGGTGCAACGGGATTGGC  
GCCTCCAGACTCTCAATGGTTCCCATGCCCTACCTACCTTGGCCCTGCATTGGGTGGAGCCACGT  
GAACACATAATCCCGTGGTGGTCAAGCCACAATGCCAGATGTGGCAGCTCCTGGTTCGAGTGAT  
CGTGTGTGCTGCAACGTGGAGGGGAGTGATGCGCAAGGTGGGCGCATGAAGGGCATGCCCA  
CGAAGCTGTGGCAGTGGGCATCCTTGTAGGCACCTGGTAGCAATAGGAATCTTCTCATCCTC  
ATTTTCACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCT  
GAAGGCGACTGTCTGAATGGCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAG  
TCCCTTGGGAGAGAGCCAGCACCCCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCTCCA  
TCTGCCCTGGGGTGGAGGCACCATCACCATCACCAGGCATGTCTGCAGAGCCTGGACACCACTT  
TATGGACTGCCCATGGGAGTGCTCCAAATGTGAGGGTGTGTTGCCCAATAATAAAGCCCCAGAGAA  
CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

**FIGURE 98**

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT  
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI  
YRRLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG  
STSLDHALERTYQLLVQVKDMGDQASGHQATATVEVSI IESTWVSLEPIHLAENLKVLYPHHMAQ  
VHWSGGDVHYHLESHPPGPFDEVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL  
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLAEDADAPGSPNSHVYQLLSPEPEDGVEGRA  
FQVDPTSGSVTLGVLPLRAGQNILLVLAMDLAGAEGGFSSTCEVEVAVTDINDHAPEFITSQIG  
PISLPEDVEPGTLVAMLTADADLEPAFRLMDFAIERGDTEGTFGLDWEPDSGHVRLRLCKNLSY  
EAAPSHEVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPISAPAGSFLLT  
IQPSDPISRTLRFSLVNDSEGWLCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTALPVPS  
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTALHWVEP  
REHIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTILVAIGIFLI  
LIFTHWTMSRKKDPDQPADSVPLKATV

**Signal peptide:**

amino acids 1-18

**Transmembrane domain:**

amino acids 762-784

## FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCACGCCTG  
 AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA  
 TCTCTTCACCTTCAAGTCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCTTGGTAC  
 ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG  
 TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCATGAGACTAGCACC  
 TCTGCCAACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG  
 TGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA  
 GCATAGTCACCAACTCTGAGTTCCATACAACTCCAGTGGGATCAGCACAGCCACCAACTCTGAG  
 TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG  
 GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCCTCCAGTGGGGCCAGCACAGTCCACCACT  
 CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC  
 AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC  
 CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA  
 CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT  
 GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG  
 AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCA  
 GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG  
 TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG  
 GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACT  
 CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCCACCAATTCTGAGTCCAGCACACCTCC  
 AGTGGGGCCAAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAAACACAGCCAC  
 CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA  
 CCTCCAGTGGGGTCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACAA  
 GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG  
 CACAGTGTCCAGTGGGATCAGCACAGTCCACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCA  
 ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA  
 ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGTGAGGCAAAGCCTGGTGGGTCCCTGGT  
 GCCGTGGGAAATCTTCTCATCACCTGGTCTCGGTTGTGGCGGCCGTGGGGCTCTTGCTGGGC  
 TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCT  
 CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCAGAGGCCAG  
 GTGGAGTCTTAAGTGGTTCTGGAGGAGACCAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA  
 ACAGCGGGCCCTGAGCAGCCCCGGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCA  
 CCAAGACCTGGTTTCTTTTCATTTCATCCAGGAGACCCCTCCAGCTTTGTTTGAGATCCTGAA  
 AATCTTGAAGAAGGTATTCTCACCTTTCTTGCTTTACCAGACACTGGAAAGAGAATACTATAT  
 TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTG  
 CCCCGGGTGGGTATCTAGCTCTGAGATGAATCAGTTATAGGAGAAAACCTCCATGCTGGACTC  
 CATCTGGCATTTCAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 100**

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS  
SVTSNGVSIIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS  
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA  
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTNGAGTATNSES  
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS  
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA  
TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESSTVSSGISTVTNSES  
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA  
AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFWRPVSII  
AMEMSGRNSGP

**Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 510-532

**FIGURE 101**

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA  
 CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCTGCTTGTCACAGGTG  
 GGAGGCTGGAACATATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCAACAAAT  
 GGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT  
 AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT  
 GGTGTCATGGAAAGGATTTACTTTTATACTGACTCTGTTTTGGGGAAGCTTTTTTGGGAAGCATTT  
 TCATGCTGAGTCCCTTTTTACCTTTGATGTTTGTAACCCATCTTGGTATCGCTGGATCAACAAC  
 CGCCTTGTTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAGT  
 GATTATAACTGGGGATGCATTTGTTCCCTGGAGAAAGAAGTGTCATTATCATGAACCATCGGACAA  
 GAATGGACTGGATGTTCCCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT  
 TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTTGGTTGGGCCATGCAGGCTGCTGCCTATAT  
 CTTCAATTCATAGGAAATGGAAGGATGACAAGAGCCATTTCGAAGACATGATTGATTACTTTTGTG  
 ATATTACGAACCACTTCAACTCCTCATATTCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG  
 TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG  
 AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG  
 ATATCACTGTGGCGTATCCTCACAACATTCCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT  
 CCCAGGGAAATCCACTTTCACGTCCACCGGTATCCAATAGACACCTCCCCACATCCAAGGAGGA  
 CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCGTTCCTTCTATCAAG  
 GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTCATTCCACCTTGCAAGTCTGAACCTCAGGGTC  
 CTTGTGGTCAAATTGCTCTCTATACTGTATTGGACCTGTTTACGCCCTGCAATGTGCCTACTCAT  
 ATATTTGTACAGTCTTGTTAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA  
 GAATATTTGGTGGACTGGAGATCATAGAACTTGCAATGTTACCGACTTTTACACAAACAGCCACAT  
 TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTG  
 GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTTCATGACTATGTGCAATATTTCTTACT  
 GCCATCATTATTTGTTAAAGATATTTTGCACTTAATTTTGTTGGGAAAAATATTGCTACAATTTT  
 TTTAATCTCTGAATGTAATTTTCGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT  
 GGGCCAGAATATTATTAAACAATCATCAGGCTTTTAAA

**FIGURE 102**

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFIMVSWKGIYFILTLFWGSFFGSI FMLS P  
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM  
FLWNCIMRYSYLRLKICLKASLKGVPFGFWAMQAAAYIFIHRKWDDKSHFEDMIDYFCDIHEP  
LQLLIFPEGTDLTENSKSRSNFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA  
YPHNIPQSEKHLQGDFFPREIH FHVHRYPIDTLPTSKEDQLWCHKRWEEKERLRSFYQGEKNF  
YFTGQSVIPPCKSELRLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG  
LEIIE LACYRL LHKQPHLNSKKNE

**Important features of the protein:**

**Signal peptide:**

amino acids 1-22

**Transmembrane domains:**

amino acids 44-63, 90-108, 354-377

**FIGURE 103**

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC  
 TCCAAATCATCCATCCACCCCTGCTGTCTCATCTGTTTTCATAGTGTGAGATCAACCCACAGGAATA  
 TCCATGGCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT  
 CACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCTCTGCTCCCTCT  
 TTCTTGAGACCAGTGCAGAGGCTATGGAAGTGCGGTTCTTCAGGAATCAGTTCCATGCTGTGGTC  
 CACCTCTACAGAGATGGGGAAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAAGTGA  
 GTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTCTAAGGCTAAAAACATCACTCCCTCGG  
 ACATCGGCCTGTATGGGTGCTGGTTTCAGTTCCAGATTACGATGAGGAGGCCACCTGGGAGCTG  
 CGGGTGGCAGCACTGGGCTCACTTCCTCTCATTTCCATCGTGGGATATGTTGACGGAGGTATCCA  
 GTTACTCTGCCTGTCTCAGGCTGGTTCCCCAGCCACAGCCAAGTGGAAGGTCCACAAGGAC  
 AGGATTTGTCTTCAGACTCCAGAGCAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC  
 ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA  
 GGTGGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTTCTA  
 TTTTACTCGGGTACTCTGTGGTGCCCTGTGTGGTGTGTCATGGGGATGATAATTGTTTTCTTC  
 AAATCCAAAGGGAAAATCCAGGCGGAAGTGGACTGGAGAAGAAAGCACGGACAGGCGAATTGAG  
 AGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCG  
 TTTCTGATCTGAAAACGTAAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGA  
 TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA  
 CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGGAAGAACA  
 ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTT  
 ACATTCAATCCCCATTTTATCAGCCTCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCCT  
 GGACTATGAGGGTGGGACCATCTCCTTCTCAATACAAATGACCAGTCCCTTATTTATACCCTGC  
 TGACATGTCAGTTTGAAGCTTGTGAGACCTTATATCCAGCATGCGATGATGACGAGGAAAAG  
 GGGACTCCCATATTTCATATGTCCAGTGTCTGGGGATGAGACAGAGAAGACCCTGCTTAAAGGGC  
 CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGTCCCGACAGGTGGCCCCAGCTTCTCT  
 CCGGAGCCTGCGCACAGAGAGTACGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCC  
 TGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCTGTGGGAG  
 TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAA  
 CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTGTAGTCACGG  
 ACAGTGATTCTGCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTGACAGGTT  
 TGAGGGCACAGTGTGCTAATGATGTGTTTTATATTATACATTTTCCACCATAAACTCTGTT  
 TGCTTATTCCACATTAATTTACTTTTTCTATACCAAATCACCCATGGAATAGTTATTGAACACC  
 TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTTCACTGATTCTATAAGCCCAGCAT  
 TACCTGATACCAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCC  
 CTCATTAAACACAGACACAAAATTTCTAAATAAAATTTTAAACAAATTAACATAAATATATTTA  
 AAGATGATATATACTACTCAGTGTGGTTGTGCCACAAATGCAGAGTTGGTTTAAATTTAAAT  
 ATCAACCAGTGAATTCAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAA

**FIGURE 104**

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH  
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR  
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGDLSSDSRANADGYSLYDVEISI  
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVMGMIIVFFK  
SKGKIQAELDWRKKGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF  
TRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT  
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG  
TPIFICPVSWG

**Signal peptide:**

amino acids 1-17

**Transmembrane domains:**

amino acids 131-150, 235-259



**FIGURE 105**

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG  
 TTTGTTGGGAACCCCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGCTCGGCAGTGTGCATTGGA  
 CTCACGTTCATTATGTGAGATATAATCAAGAAGACCTACAATTACTATAGCACATTGTCATTTACAAC  
 TGACAAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAATTTTACAGAAATGAGCCAGAGACTTGAAT  
 CAATGGTGAAAAATGCATTTTATAAATCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC  
 AGTCAACAGAAGCATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTCTACTCTACTGAGGATCCTGA  
 AACTGTAGATAAAATTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG  
 ATCCTCACTCAGTTAAATTAATAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGA  
 ACACGAAGAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATG  
 GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAATGCCACATGGC  
 TTGTGAGTGTGCTCACTGTTTACAACATATAAGAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACA  
 ATAAACCTTCGAAATGAAACGGGGTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACA  
 TGACTATGATATTTCTCTTGACAGACTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCT  
 TCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTGTGACAGGATTTGGAGCACTGAAAAAT  
 GATGGTTACAGTCAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC  
 TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT  
 GCCAGGGTGACTCTGGAGGACCACTGGTTAGTTCAGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTG  
 AGCTGGGGAGATGAATGTGCGAAACCAACAAGCCTGGTGTATTAAGTACAGTTACGGCCTTGCGGGACTG  
 GATTACTTCAAAAACCTGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTGTTTTTTG  
 GGTGTGGAGGCCATTTTLAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTGACTGATCTCA  
 ATAAACTGTTTGCTTGATGCATGTATTTCTTCCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCA  
 GATCAACTCTGTCTGTGAGCAATAGTTGAAACTTTATGTACATAGAGAAATAGATAATACAATATTAC  
 ATTACAGCCTGTATTCAATTGTTCTCTAGAAGTTTGTGAGAATTTGACTTGTGACATAAATTTGTAAT  
 GCATATATACAATTTGAAGCACTCCTTTCTTCAGTTTCTCAGCTCCTCTCATTTCAGCAAATATCCATTT  
 TCAAGGTGCAGAACAGGAGTGAAAGAAAATATAAGAAGAAAAAATCCCTACATTTTATTTGGCACAGAA  
 AAGTATTAGGTGTTTTCTTAGTGAATATTAGAAATGATCATATTCAATTATGAAAGGTCAAGCAAAGACA  
 GCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAACCTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG  
 ATATATCCTTATTTTCAATTTCCAAACAACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTGTGACCT  
 ATAATAATTATACAACTTCATGCAATGTACTTGTCTAAGCAAATTAAGCAAATATTTATTTAACATTG  
 TTACTGAGGATGTCAACATATAACAATAAATATAAATCACCCA

## **FIGURE 106**

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTNYNSTLSFTTDKLY  
 AEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED  
 PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG  
 TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL  
 RRIIVHEKYKHPSHDYDISLAELSSPVPYTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS  
 QNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSSGGLVSSDARDIWYLAG  
 IVSWGDECAKPNKPGVYTRVTALRDWITSKGTGI

**Transmembrane domain:**

amino acids 21-40 (type II)

**FIGURE 107**

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCTG  
 CCCCAGATGAGCCCCCGCGTGCCTCCCGACTATCCCCAGGCGGGCGTGGGGCACCGGGCCCAGC  
 GCCGACGATCGCTGCCGTTTTGCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT  
 ACGGGGCTCACAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTACGCCCTCAA  
 TCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGATGAGGGACTACCTAA  
 ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTTGACTTACTTTCTCT  
 GTGGTTCATCCGGTCATGATTGCTGTTTGGCTGTTTCCTTATCATTGTGGGGATGTTAGGATATTG  
 TGGAACGGTGAAAAGAAATCTGTTGCTTCTTGCATGGTACTTTGGAAGTTTGCTTGTCATTTTCT  
 GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTTATGGTTCAGTACAATGGTCA  
 GATATGGTCACCTTTGAAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA  
 TGCTTGGAAATTTTTTTCAGAGAGAGTTAAGTGCTGTGGAGTAGTATATTTCACTGACTGGTTGG  
 AAATGACAGAGATGGACTGGCCCCAGATTCTGCTGTGTAGAGAATTTCCAGGATGTTCCAAA  
 CAGGCCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTGTGGGAAGAAATGTATTCCTT  
 TTTGAGAGGAACCAACAACCTGCAGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAA  
 TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG  
 ACAGACCAATGATGTCTTGAAGAATGACAACTCTCAGCACCTGTATGTCCCTCAGTAGAACT  
 GTTGAAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACT  
 TTGAGATGGAGGAGTTATATAAAGAAATGTCACAGAAGAAAACCACAACTTGTTTTATTGGACT  
 TGTGAATTTTTGAGTACATACTATGTGTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAA  
 TAACACCTAAGCATATACTATTCTATGCTTTAAATGAGGATGGAAAAGTTTCATGTCATAAGTC  
 ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC  
 TGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCCGCA  
 TCCATGCAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTTACTTCTACCAA  
 CTAGTATATAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA  
 CTCAGCGATCTATTCTTCTGATGCTAAATAAATTATATATCAGAAAACCTTTCAATATTTGGTGACT  
 ACCTAAATGTGATTTTTGCTGGTTACTAAAATATTCTTACCACCTAAAAGAGCAAGCTAACACAT  
 TGTCTTAAGCTGATCAGGGATTTTTTGTATATAAGTCTGTGTTAAATCTGTATAATTCAGTCGAT  
 TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCCTGTATAGCATCATT  
 ATTTTTAGCCTTTCTGTTAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAAC  
 TGTATTATTAATACTTAACCACTAATTTTGAAAATTACCAGTGTGATACATAGGAATCATTATTC  
 AGAATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTGAGA  
 AAGGACTTGTATGCTGTTTTTCTCCAAATGAAGACTCTTTTGACACTAAACACTTTTAAAAA  
 GCTTATCTTTGCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATTTACAGAAAA  
 TAGTGTTCTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAAACATGTGACAATTTAGAGATT  
 CTTTGTATTTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAATTTTTTTACAA  
 GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCATTTTACTGTATTTTGTGTATTTTGTAT  
 TTCTCAGAATATGGAAGAAAATTAATGTGTCAATAAATATTTTCTAGAGAGTAA

## **FIGURE 108**

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVVHP  
 VMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWVTEQELMVPVQWSDMVT  
 LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFPGCSKQAHQ  
 EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPQDQM  
 MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

**Signal peptide:**

amino acids 1-33

**Transmembrane domains:**

amino acids 12-35, 57-86, 94-114, 226-248

**FIGURE 109**

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCTTCCTCTGATAAAGCCCCCTACCAGTGCT  
 GATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAAAGTGCTTGAAAGAGAAGGGGACAAAGGAACA  
 CCAGTATTAAGAGGATTTCCAGTGTTTCTGGCAGTTGGTCCAGAAGGATGCTCCATTCTGCTTCTCACCTG  
 CCTCTTCATCACAGGCACCTCCGTGTACCCGTGGCCCTAGATCCTTGTCTGCTTACATCAGCCTGAATGAGC  
 CCTGGAGGAACACTGACCACCAGTTGGATGAGTCTCAAGGTCCTCCTCTATGTGACAACCATGTGAATGGGAG  
 TGGTACCACCTTCACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA  
 CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAAGGCACGGCATTGTGCAACGCCAGGCTTGTGCCAGCT  
 TCAATGGGAACGTCTGTCTCTGGAACACCACGGTGGAGTCAAGGCTTGCCTGGAGGCTACTATGTGTATCGT  
 CTGACCAAGCCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTTATGACATCTGCGACGAGGACTGCCATGG  
 CAGCTGCTCAGATACCAGCGAGTGCACATGCCCTCCAGGAACGTGTCTAGGCCCTGACAGGCAGACATGCTTTG  
 ATGAAAATGAATGTGAGCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTCTACCGCTGT  
 GAGTGTGGGGTTGGCCGTGTGCTAAGAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA  
 TGGTGGCTGCAGCCACTCTTGCCTTGGATCTGAGAAAGGCTACCACTGTGAATGTCCCCGGGGCTGGTGTGT  
 CTGAGGATAACCACACTTGCCAAGTCCCTGTGTGTGCAAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG  
 CTGGTTGGTGGCCTGGAGCTTCTCTGACCAACACCTCCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACAT  
 CCTCTTCTCTCTCAAGACATGTGGTACAGTGGTCGATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA  
 CAGGTCTACCCAAGCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACAGCAAGCTGCTGATCCCGGTG  
 ACCTGCGAGTTTCCACGCCTGTACACCATTTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCCACTGGAAAT  
 CATGAGCCGAAATCATGGGATCTTCCCATTCACTCTGGAGATCTTCAAGGACAATGAGTTTGAAGAGCCTTACC  
 GGGAAAGCTCTGCCCACCCTCAAGCTTCTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC  
 TTGAAAAGCTTGGTGGAGAGCTGCTTTGCCACCCCCACCTCCAAGATCGACGAGGTCTGAAATACTACCTCAT  
 CCGGGATGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGG  
 TCCCTGTCTTCAAGTTTGTGGGCAAGACCACAAGGAAGTGTCTGCACTGCCGGGTTCTTGTCTGTGGAGTG  
 TTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGGCAATGCGTCGTGGGGCAGGAGGAGAGGACTCAGC  
 CGGTCTACAGGGCCAGACGCTAACAGGCGGCCCGATCCGCATCGACTGGGAGGACTAGTTTCGTAGCCATACCTC  
 GAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCCACC GGCCCTTAAGAACATCTGCCAACAGC  
 TGGGTTTCAGACTTCACACTGTGAGTTTCAGACTCCCAGCACCAACTCACTCTGATTCTGGTCCATTCACTGGGCA  
 CAGGTCACAGCACTGCTGAACAATGTGGCCTGGGTGGGGTTTCATCTTTCTAGGGTTGAAAACCTAACTGTCCA  
 CCCAGAAAGACACTCACCCATTTCCTCATTCTTTCTTCTACACTTAAATACCTCGTGTATGGTGCAATCAGAC  
 CACAAAATCAGAAGCTGGGTATAATATTCAAGTTACAAACCCTAGAAAAATTAACAGTTACTGAAATTATGA  
 CTTAAATACCAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAATTTCAATTCAAATGCAGACTAA  
 TTATAGGGAATTTGAAGTGTATCAATAAACAGTATATAATTTT

2

**Important features of the protein:**

amino acids 1-16

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

amino acids 411-418, 443-451

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,  
522-528, 531-537

## amino acids 197-209

## amino acids 431-457

## amino acids 191-212, 232-253

**FIGURE 111**

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG  
 GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGAGGCCCT  
 CCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGAGGATCACAGAGCCAGCATGT  
 TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCC  
 ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATATAGCACTACTGAGCCTGGCGAGTATCATCATTGT  
 GGTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGA  
 GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC  
 CCCGAAGGGCCTGCAGTGGCAGTCCGCCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC  
 AGGGAACCTGGTTCTCTGCCTGTTTCGACAACCTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG  
 GCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGATGTTGTGTGAAATCACAGAAAACGCCAG  
 GAGCTTCGATGCGGAACCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCCTGTCTTGCCTG  
 TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGTGGGAGGAGGCCCTCTGTGATTCTTGGCCTTGGCAGG  
 TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCA  
 GCCCCTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGGCAG  
 CTTCCCATCCCTGGCTGTGGCCAAGATCATCATATTGAATTCAACCCCATGTACCCCAAAGACAATGACA  
 TCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTT  
 GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTTGGATGGGGCTTACGAAGCAGAATGGAGGAA  
 GATGTCTGACATACTGTGTCAGGCGTCAGTCCAGGTCATTGACAGCACACGGTGCAATGCAGACGATGCGT  
 ACCAGGGGGAAGTCACCGAGAAGATGATGTGTGAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGT  
 GACAGTGGTGGGCCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG  
 CTGCGGGGGCCCCGAGCACCCCAAGGATATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT  
 GGAAGGCTGAGCTGTAATGCTGCTGCCCTTTGCAGTGTCTGGGAGCCGCTTCCTTCCTGCCCTGCCACCT  
 GGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCCTGGGTACACCCCTCTGCCCACAGCCTCAGCAT  
 TTCTTGAGCAGCAAAGGGCCTCAATTCTGTAAAGAGACCTCGCAGCCCAGAGGCGCCCAGAGGAAGTCA  
 GCAGCCCTAGCTCGGCCACACTTGGTGCTCCCAGCATCCAGGGAGAGACACAGCCCACTGAACAAGGTCT  
 CAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAGGCC  
 CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGCTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAA  
 GCCTACTAGAGCAAGAAACAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT  
 ACTGTTGTATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA  
 AAAA

## **FIGURE 112**

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIIALLSLASIIIVVLIKVILDKYYFLCG  
 QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN  
 FTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL  
 KTPRVVGEEASVDSWPWQVSIQYDKQHVCSSILDPHWVLTAAHCFRKHTDVFNWKVRAGSDKL  
 GSFPSSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIGWG  
 FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCQGDSGGPLMYQS  
 DQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

**Transmembrane domain:**

amino acids 32-53 (typeII)



# **FIGURE 113**

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA  
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACTTTTTCAGCAACTAAAAAGCCACAGGAGT  
TGAAGTGTCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAAAATC  
TGTTTTTTGTTCTCTTGTAAGTACGCTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT  
GGCCCAAACCTGACCTTCACCTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCCTCGAAG  
CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAACTGTCTCACGTCT  
GGAGGCACTGACTCGGGCAGTGACAGGTAGCTGAGCCTCTGGTAGCTGCGGCTTTCAAGGTGGGC  
CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTTTCATAGGCGATGGCTCCCACTGCCC  
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGCCGTGGACACCTGCTCA  
GAAGCAGTGGGTGAGACATCACGCTGCCCCGCCATCTAACCTTTTCATGTCTGCACATCACCTG  
ATCCATGGGCTAATCTGAACTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC  
CAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGACTTGCACTCTCTGGAACATGAGG  
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACCTGTGCCAAATTATGGGTGAGAAAAGATG  
GAGGTGTTGGGTATCACAAAGGCATCGAGTCTCCTGCATTCAGTGACATGTGGGGGAAGGGCTG  
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGAT  
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA  
TCTGCGATCACCAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCC  
CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG  
GGCTAATGGCTCAGTGTGGCCAGGAGGTCAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT  
GTGCGAACACGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGG  
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCAAGACCAAAGGAGCTAGAGCTTGTT  
CAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAAATCCA  
AACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT  
TTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGG  
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC  
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACTGGGAGGTGAAGGAGGCTGAGACA  
GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAATAAAAAAGAATTA  
TGGTTATTTGTAA

# **FIGURE 114**

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC  
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

**Signal peptide:**  
amino acids 1-15

**FIGURE 115**

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA  
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA  
ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAATTGTCCT  
GTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAAGCCTATGACATGGAGCACACTT  
TCTACAGCAATGGAGAGAAGAAGAAGATTACATGGAAATTGATCCTGTGACCAGAAGTGAATA  
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG  
CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAAACTCAGATTAAAGTGATTCTGAATTTT  
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACTTCTTTGAACAGTCAGTG  
ATTTGGGTCCCAGCAGAAAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCCAAAATCTGGA  
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG  
ACTTTGAGGAGGAGGAGAAGATCTTCACTTTCTGCCAACGAAAAAAAGGGATTGAACAAAAT  
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA  
AGAACTTCCAATAAATGACTATACTGAAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAG  
GTTATTGTTGATTTTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCTCTGTGAACCTTTACTA  
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTG  
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTTAATAGGAGGTTTGAGCTCAAATGCTTAAAC  
TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCATGAGGCATCTGGCCCT  
GGTAGCCAGCTCTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACA  
TTATCACCAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 116**

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPPKAYDME  
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKTQIKVIP  
EFSEPEEEIDENEEITTTFFEQSVIWPVPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVSE  
LQDFEEEGEDLHFPANEKKGIEQNEQWVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD  
ERGYCCIIYCRRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-40

#### **Transmembrane domain:**

amino acids 25-47 (type II)

#### **N-glycosylation sites.**

amino acids 94-97, 180-183

#### **Glycosaminoglycan attachment sites.**

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-  
242

#### **N-myristoylation sites.**

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

#### **Microbodies C-terminal targeting signal.**

amino acids 315-317

#### **Cytochrome c family heme-binding site signature.**

amino acids 9-14

**FIGURE 117**

GAGCTCCCCCTCAGGAGCGGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGGCGGCA  
 GGGCGGGCGGCCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCGTTCCCTCTGTCCATCCTGGGGCT  
 GGCCGGCTGCATCGCGGCCACC GGATGGACATGTGGAGCAGGACCTGTACGACAACCCCGTCACCT  
 CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCC  
 TATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCGAGCCCTGATGATCGTAGGCATCGTCTT  
 GGGTGCCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG  
 CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTTCATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTG  
 TCTGTGTTTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG  
 GATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGGCGCTCTGTTCTGTGGGCTGGGTCTGGAGGCC  
 TCACACTAATTGGGGGTGTGATGATGTGCATCGCTGCCGGGGCTGGCACCAGAAGAAACCAACTACAAA  
 GCCGTTTCTTATCATGCCTCAGGCCACAGTGTTCCTTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTT  
 TGGGTCCAACACCAAAAACAAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATC  
 CTTCCAAGCAGCACTATGTGTAATGTCTTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCA  
 CCAAAAAACAAGGAGATCCCCTCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCT  
 CGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCACCATAAAACA  
 GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTTTAAATATAACT  
 TTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTC  
 TTCCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAAAACCTTTTGAAAGGAAA  
 GAGTAGACCCAAAGATGTTATTTTCTGCTGTTTGAATTTTGTCTCCCCACCCCAACTTGCGTAGTAATAA  
 ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTT  
 ACACTGTGATCTTAAAAGTTACCAAACCAAAGTCATTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTG  
 TTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCCTGAGCTCTCCACTGGAGTCCTCTTTCTGT  
 CGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTAATTTAAGTCCTAAATATAGTTAA  
 AATAAATAATGTTTTAGTAAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG  
 GAAATGAAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTC  
 ATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGT  
 TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCA  
 TGGTGGCATAACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGT  
 TGGGGCTGCAGTGAGCCATGATCACCACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA  
 AATAAAAAATAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAACTAATTCTTTAA

## **FIGURE 118**

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP  
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL  
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA  
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY  
V

**Signal peptide:**

amino acids 1-23

**Transmembrane domains:**

amino acids 81-100, 121-141, 173-194

**FIGURE 119**

GGAAAACTGTTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCGGAGTCC  
 AGCTGGCTAAAACTCATCCCAGAGGATAAATGGCAACCCATGCCTTAGAAATCGCTGGGCTGTTCTTG  
 GTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGGAGAGTGTGGCCTTCATT  
 GAAAACAACATCGTGGTTTTTGAAAACCTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA  
 CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG  
 GACTGATGTGTGCTGCTTCCGTGATGTCCTTCTTGCTTTTCATGATGGCCATCCTTGGCATGAAATGC  
 ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT  
 CATCACGGGCATGGTGGTGTCTATCCCTGTGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTATA  
 ACTCAATAGTGAATGTTGCCCAAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCA  
 CTGGTGTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTA  
 CAGATACTCGATACCTTCCCATCGCACAAACCAAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCG  
 TCTACTCCAGAAGTCAGTATGTGTAGTTGTGTATGTTTTTTAACTTTACTATAAAGCCATGCAAATG  
 ACAAAAATCTATATTACTTTCTCAAAATGGACCCAAAGAACTTTGATTACTGTTCTTAACTGCCT  
 AATCTTAATTACAGGAACTGTGCATCAGCTATTTATGATTCTATAAGCTATTTTCAGCAGAATGAGATA  
 TTAAACCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA  
 CTCTTTTTATCATTACTTCAAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCC  
 ACGACATAGCATTATGTACATAGATGAGTGTAACATTTATATCTCACATAGAGACATGCTTATATGGT  
 TTTATTTAAAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAA  
 ATCATGGATAGGGTTGAAGAAGGTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCCTCCA  
 TTTATAATGAAGATTAAAATGAAGGCTTAAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATAT  
 GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCAGAGGCTTTTTTTTT  
 CTTGTGTATTAAATTAACATTTTAAAAACGCAGATATTTGTCAAGGGGCTTGCATTCAAACCTGCTT  
 TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAAAGTG  
 AAAATATTTTTGTTTTGTATTGTAAGAAGAATGATGCATTTTGACAAGAAATCATATATGTATGGAT  
 ATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAAGAGCAGAAAAATA  
 TGTCTTGGTTTTTCATTGCTTACCAAAAAACAACAACAAAAAAGTTGCCTTTGAGAACTTCACCT  
 GCTCCTATGTGGGTACCTGAGTCAAAATGTCAATTTTGTCTGTGAAAAATAAATTTCTTCTGTGA  
 CCATTTCTGTTTAGTTTTACTAAAATCTGTAAATACTGTATTTTCTGTTTATTCCAAATTTGATGAA  
 ACTGACAATCCAATTTGAAAGTTTGTGTGACGCTGTCTAGCTTAAATGAATGTGTTCTATTGCTT  
 TATACATTTATATTAATAAATGTACATTTTCTAATT

**FIGURE 120**

MATHALEIAGLFLGGVMVGTVAVTVMPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRMQCK  
IYDSLALALSPDLQAARGLMCAASVMSFLAFMAILGMKTRCTGDNEKVKAHILLTAGIIFIITG  
MVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY  
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

**Signal peptide:**

amino acids 1-17

**Transmembrane domains:**

amino acids 82-101, 118-145, 164-188





## **FIGURE 122**

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAQVPGR  
 DGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK  
 MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGGLPIEAIYLDQGSPEMNSTINIHRIS  
 SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIIIEELPK

**Signal peptide:**

amino acids 1-30

**Transmembrane domain:**

amino acids 195-217

**FIGURE 123**

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA  
 CTGGGTGCTCATCACGGGAACCTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA  
 ATTGCCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTTTTTA  
 ACCGCCCCCTCCCCACCCCCCAAAAAAAGCTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC  
 TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGGAG  
 TGTTCCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAGGGGTCCAATTTT  
 TCTTCCTGGGTGTGAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTGTCATGCAACTG  
 GCCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAACAATACAAAGGATGGGGTTTCAATG  
 TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCCACTGTCTTACTGACAATG  
 CTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTGA  
 ATCTCAGAAATTACAGGAGATACCCCTCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT  
 ATAACAGCCTTCAAAAACTTAAGTATAATCAATTTAAAGGGCTCAACCAGCTCACCTGGCTATAC  
 CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTTTAAATGGAATACGCAGACTCAAAGA  
 GCTGATTCTTAGTTCCAATAGAATCTCCTATTTTCTTAACAATACCTTCAGACCTGTGACAAAAT  
 TACGGAACCTGGATCTGTCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTG  
 CGGAAGCTGCTGAGTTTACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA  
 AGACTGCCGCAACCTGGAACTTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG  
 TCTTTGCTGGCATGATCAGACTCAAAGAACTTCACCTGGAGCACAATCAATTTTCCAAGCTCAAC  
 CTGGCCCTTTTCCAAGGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGGAATAAAATCAGTGT  
 CATAGGACAGACCATGTCTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGA  
 TCGAAGCTTTCAGTGGACCCAGTGTTTTCCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGAT  
 TCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAG  
 TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTGCTCCCTTGTAACCTGGCTGAAAAGTT  
 TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG  
 ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTTGTCTGGCCAG  
 GGCTCTCCCAAAGCCGACGTTTAAAGCCCAAGCTCCCGAGGCCGAAGCATGAGAGCAAACCCCTT  
 TGCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCT  
 TTCCATAAAATCATCGCGGCAGCGTGCGCTTTTCTGTCCGTGCTCGTCATCTGCTGGTTAT  
 CTACGTGTCTATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA  
 GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCCAGCACCCAGGAATTTTATGTA  
 GATTATAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCTGCACCTA  
 TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTGTGATAAAAAGAGCTCTTAAAGCT  
 GGGAAATAAGTGGTGCTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCCTCCCC  
 TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCCTGTCCGTTTGTAGTGATTATAATACT  
 GGTCATTTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT  
 GAACTCCGTTTAAATATAATACCTATTGTATAAGACCCTTTACTGATTCCATTAATGTGCGATTT  
 GTTTTAAGATAAACTTCTTTCATAGGTAAAAA

## FIGURE 124

MGFNVIRLLSGSAVALVIAPTPLLTLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGCLG  
LSLRYSLSQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR  
PVTNLRNLDSLQNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRIQDCRNLELLDLGYNRIRS  
LARNVFAGMIRLKEHLHLEHNQFSKLNALFPRLVSLQNLQWKNISVIGQTMSTWSSQLRDL  
SGNEIEAFSGPSVFQCVPNLQRLNLDNKLTFIGQEILDSWISLNDISLAGNIWECRNICSLVN  
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE  
SKPPLPPTVGATEPGPETDADAHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMQLQQR  
SLMRRHRKKKRQSLKQMPSTQEFYVDYKPTNTETSEMLLNGTGPCITYNKSGSRECEV

### Important features of the protein:

#### Signal peptide:

amino acids 1-33

#### Transmembrane domain:

amino acids 420-442

#### N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

#### cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

#### Tyrosine kinase phosphorylation site.

amino acids 136-142

#### N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

1. NAME \_\_\_\_\_  
 2. ADDRESS \_\_\_\_\_  
 3. CITY \_\_\_\_\_  
 4. STATE \_\_\_\_\_  
 5. ZIP \_\_\_\_\_  
 6. PHONE \_\_\_\_\_  
 7. DATE \_\_\_\_\_  
 8. SIGNATURE \_\_\_\_\_  
 9. PRINT NAME \_\_\_\_\_  
 10. PRINT ADDRESS \_\_\_\_\_  
 11. PRINT CITY \_\_\_\_\_  
 12. PRINT STATE \_\_\_\_\_  
 13. PRINT ZIP \_\_\_\_\_  
 14. PRINT PHONE \_\_\_\_\_  
 15. PRINT DATE \_\_\_\_\_  
 16. PRINT SIGNATURE \_\_\_\_\_  
 17. PRINT NAME \_\_\_\_\_  
 18. PRINT ADDRESS \_\_\_\_\_  
 19. PRINT CITY \_\_\_\_\_  
 20. PRINT STATE \_\_\_\_\_  
 21. PRINT ZIP \_\_\_\_\_  
 22. PRINT PHONE \_\_\_\_\_  
 23. PRINT DATE \_\_\_\_\_  
 24. PRINT SIGNATURE \_\_\_\_\_  
 25. PRINT NAME \_\_\_\_\_  
 26. PRINT ADDRESS \_\_\_\_\_  
 27. PRINT CITY \_\_\_\_\_  
 28. PRINT STATE \_\_\_\_\_  
 29. PRINT ZIP \_\_\_\_\_  
 30. PRINT PHONE \_\_\_\_\_  
 31. PRINT DATE \_\_\_\_\_  
 32. PRINT SIGNATURE \_\_\_\_\_  
 33. PRINT NAME \_\_\_\_\_  
 34. PRINT ADDRESS \_\_\_\_\_  
 35. PRINT CITY \_\_\_\_\_  
 36. PRINT STATE \_\_\_\_\_  
 37. PRINT ZIP \_\_\_\_\_  
 38. PRINT PHONE \_\_\_\_\_  
 39. PRINT DATE \_\_\_\_\_  
 40. PRINT SIGNATURE \_\_\_\_\_  
 41. PRINT NAME \_\_\_\_\_  
 42. PRINT ADDRESS \_\_\_\_\_  
 43. PRINT CITY \_\_\_\_\_  
 44. PRINT STATE \_\_\_\_\_  
 45. PRINT ZIP \_\_\_\_\_  
 46. PRINT PHONE \_\_\_\_\_  
 47. PRINT DATE \_\_\_\_\_  
 48. PRINT SIGNATURE \_\_\_\_\_  
 49. PRINT NAME \_\_\_\_\_  
 50. PRINT ADDRESS \_\_\_\_\_  
 51. PRINT CITY \_\_\_\_\_  
 52. PRINT STATE \_\_\_\_\_  
 53. PRINT ZIP \_\_\_\_\_  
 54. PRINT PHONE \_\_\_\_\_  
 55. PRINT DATE \_\_\_\_\_  
 56. PRINT SIGNATURE \_\_\_\_\_  
 57. PRINT NAME \_\_\_\_\_  
 58. PRINT ADDRESS \_\_\_\_\_  
 59. PRINT CITY \_\_\_\_\_  
 60. PRINT STATE \_\_\_\_\_  
 61. PRINT ZIP \_\_\_\_\_  
 62. PRINT PHONE \_\_\_\_\_  
 63. PRINT DATE \_\_\_\_\_  
 64. PRINT SIGNATURE \_\_\_\_\_  
 65. PRINT NAME \_\_\_\_\_  
 66. PRINT ADDRESS \_\_\_\_\_  
 67. PRINT CITY \_\_\_\_\_  
 68. PRINT STATE \_\_\_\_\_  
 69. PRINT ZIP \_\_\_\_\_  
 70. PRINT PHONE \_\_\_\_\_  
 71. PRINT DATE \_\_\_\_\_  
 72. PRINT SIGNATURE \_\_\_\_\_  
 73. PRINT NAME \_\_\_\_\_  
 74. PRINT ADDRESS \_\_\_\_\_  
 75. PRINT CITY \_\_\_\_\_  
 76. PRINT STATE \_\_\_\_\_  
 77. PRINT ZIP \_\_\_\_\_  
 78. PRINT PHONE \_\_\_\_\_  
 79. PRINT DATE \_\_\_\_\_  
 80. PRINT SIGNATURE \_\_\_\_\_  
 81. PRINT NAME \_\_\_\_\_  
 82. PRINT ADDRESS \_\_\_\_\_  
 83. PRINT CITY \_\_\_\_\_  
 84. PRINT STATE \_\_\_\_\_  
 85. PRINT ZIP \_\_\_\_\_  
 86. PRINT PHONE \_\_\_\_\_  
 87. PRINT DATE \_\_\_\_\_  
 88. PRINT SIGNATURE \_\_\_\_\_  
 89. PRINT NAME \_\_\_\_\_  
 90. PRINT ADDRESS \_\_\_\_\_  
 91. PRINT CITY \_\_\_\_\_  
 92. PRINT STATE \_\_\_\_\_  
 93. PRINT ZIP \_\_\_\_\_  
 94. PRINT PHONE \_\_\_\_\_  
 95. PRINT DATE \_\_\_\_\_  
 96. PRINT SIGNATURE \_\_\_\_\_  
 97. PRINT NAME \_\_\_\_\_  
 98. PRINT ADDRESS \_\_\_\_\_  
 99. PRINT CITY \_\_\_\_\_  
 100. PRINT STATE \_\_\_\_\_  
 101. PRINT ZIP \_\_\_\_\_  
 102. PRINT PHONE \_\_\_\_\_  
 103. PRINT DATE \_\_\_\_\_  
 104. PRINT SIGNATURE \_\_\_\_\_  
 105. PRINT NAME \_\_\_\_\_  
 106. PRINT ADDRESS \_\_\_\_\_  
 107. PRINT CITY \_\_\_\_\_  
 108. PRINT STATE \_\_\_\_\_  
 109. PRINT ZIP \_\_\_\_\_  
 110. PRINT PHONE \_\_\_\_\_  
 111. PRINT DATE \_\_\_\_\_  
 112. PRINT SIGNATURE \_\_\_\_\_  
 113. PRINT NAME \_\_\_\_\_  
 114. PRINT ADDRESS \_\_\_\_\_  
 115. PRINT CITY \_\_\_\_\_  
 116. PRINT STATE \_\_\_\_\_  
 117. PRINT ZIP \_\_\_\_\_  
 118. PRINT PHONE \_\_\_\_\_  
 119. PRINT DATE \_\_\_\_\_  
 120. PRINT SIGNATURE \_\_\_\_\_  
 121. PRINT NAME \_\_\_\_\_  
 122. PRINT ADDRESS \_\_\_\_\_  
 123. PRINT CITY \_\_\_\_\_  
 124. PRINT STATE \_\_\_\_\_  
 125. PRINT ZIP \_\_\_\_\_  
 126. PRINT PHONE \_\_\_\_\_  
 127. PRINT DATE \_\_\_\_\_  
 128. PRINT SIGNATURE \_\_\_\_\_  
 129. PRINT NAME \_\_\_\_\_  
 130. PRINT ADDRESS \_\_\_\_\_  
 131. PRINT CITY \_\_\_\_\_  
 132. PRINT STATE \_\_\_\_\_  
 133. PRINT ZIP \_\_\_\_\_  
 134. PRINT PHONE \_\_\_\_\_  
 135. PRINT DATE \_\_\_\_\_  
 136. PRINT SIGNATURE \_\_\_\_\_  
 137. PRINT NAME \_\_\_\_\_  
 138. PRINT ADDRESS \_\_\_\_\_  
 139. PRINT CITY \_\_\_\_\_  
 140. PRINT STATE \_\_\_\_\_  
 141. PRINT ZIP \_\_\_\_\_  
 142. PRINT PHONE \_\_\_\_\_  
 143. PRINT DATE \_\_\_\_\_  
 144. PRINT SIGNATURE \_\_\_\_\_  
 145. PRINT NAME \_\_\_\_\_  
 146. PRINT ADDRESS \_\_\_\_\_  
 147. PRINT CITY \_\_\_\_\_  
 148. PRINT STATE \_\_\_\_\_  
 149. PRINT ZIP \_\_\_\_\_  
 150. PRINT PHONE \_\_\_\_\_  
 151. PRINT DATE \_\_\_\_\_  
 152. PRINT SIGNATURE \_\_\_\_\_  
 153. PRINT NAME \_\_\_\_\_  
 154. PRINT ADDRESS \_\_\_\_\_  
 155. PRINT CITY \_\_\_\_\_  
 156. PRINT STATE \_\_\_\_\_  
 157. PRINT ZIP \_\_\_\_\_  
 158. PRINT PHONE \_\_\_\_\_  
 159. PRINT DATE \_\_\_\_\_  
 160. PRINT SIGNATURE \_\_\_\_\_  
 161. PRINT NAME \_\_\_\_\_  
 162. PRINT ADDRESS \_\_\_\_\_  
 163. PRINT CITY \_\_\_\_\_  
 164. PRINT STATE \_\_\_\_\_  
 165. PRINT ZIP \_\_\_\_\_  
 166. PRINT PHONE \_\_\_\_\_  
 167. PRINT DATE \_\_\_\_\_  
 168. PRINT SIGNATURE \_\_\_\_\_  
 169. PRINT NAME \_\_\_\_\_  
 170. PRINT ADDRESS \_\_\_\_\_  
 171. PRINT CITY \_\_\_\_\_  
 172. PRINT STATE \_\_\_\_\_  
 173. PRINT ZIP \_\_\_\_\_  
 174. PRINT PHONE \_\_\_\_\_  
 175. PRINT DATE \_\_\_\_\_  
 176. PRINT SIGNATURE \_\_\_\_\_  
 177. PRINT NAME \_\_\_\_\_  
 178. PRINT ADDRESS \_\_\_\_\_  
 179. PRINT CITY \_\_\_\_\_  
 180. PRINT STATE \_\_\_\_\_  
 181. PRINT ZIP \_\_\_\_\_  
 182. PRINT PHONE \_\_\_\_\_  
 183. PRINT DATE \_\_\_\_\_  
 184. PRINT SIGNATURE \_\_\_\_\_  
 185. PRINT NAME \_\_\_\_\_  
 186. PRINT ADDRESS \_\_\_\_\_  
 187. PRINT CITY \_\_\_\_\_  
 188. PRINT STATE \_\_\_\_\_  
 189. PRINT ZIP \_\_\_\_\_  
 190. PRINT PHONE \_\_\_\_\_  
 191. PRINT DATE \_\_\_\_\_  
 192. PRINT SIGNATURE \_\_\_\_\_  
 193. PRINT NAME \_\_\_\_\_  
 194. PRINT ADDRESS \_\_\_\_\_  
 195. PRINT CITY \_\_\_\_\_  
 196. PRINT STATE \_\_\_\_\_  
 197. PRINT ZIP \_\_\_\_\_  
 198. PRINT PHONE \_\_\_\_\_  
 199. PRINT DATE \_\_\_\_\_  
 200. PRINT SIGNATURE \_\_\_\_\_  
 201. PRINT NAME \_\_\_\_\_  
 202. PRINT ADDRESS \_\_\_\_\_  
 203. PRINT CITY \_\_\_\_\_  
 204. PRINT STATE \_\_\_\_\_  
 205. PRINT ZIP \_\_\_\_\_  
 206. PRINT PHONE \_\_\_\_\_  
 207. PRINT DATE \_\_\_\_\_  
 208. PRINT SIGNATURE \_\_\_\_\_  
 209. PRINT NAME \_\_\_\_\_  
 210. PRINT ADDRESS \_\_\_\_\_  
 211. PRINT CITY \_\_\_\_\_  
 212. PRINT STATE \_\_\_\_\_  
 213. PRINT ZIP \_\_\_\_\_  
 214. PRINT PHONE \_\_\_\_\_  
 215. PRINT DATE \_\_\_\_\_  
 216. PRINT SIGNATURE \_\_\_\_\_  
 217. PRINT NAME \_\_\_\_\_  
 218. PRINT ADDRESS \_\_\_\_\_  
 219. PRINT CITY \_\_\_\_\_  
 220. PRINT STATE \_\_\_\_\_

APP\_ID=10063524

**FIGURE 126**

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGARES  
APYRGMVRTALGIIIEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVGKSEDEHYPLWKS  
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP  
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIMNQP  
RDKQGRGLLYKSSTDCLIQAVQGEFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

**Transmembrane domains:**

amino acids 25-38, 130-147, 233-248

**FIGURE 127**

CGCGGATCGGACCCAAGCAGGTTCGGCGGCGGGCGGCAGGAGAGCGGCCGGGCGTCAGCTCCTCGAC  
CCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGCCAGGCCCGGCATGG  
AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCGCTCGGGGGGCATCGGCGCGGCCGTGGCC  
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA  
GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACTTTGATCCCCCTACAGATGTGACCTAT  
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC  
TGCATCAACAATGCTGGCTTGGCCCGGCGCTGACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAA  
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA  
AGGAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTTA  
CCCCGTGCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACT  
GAGGCAAGAGCTTCGGGAGGCCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG  
AGACACAATTGCGCTTCAAACCTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA  
ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCCTCAGCACCCCCGCACA  
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTCC  
TCCTTCCCTCCCCACCCCTTCATGGCTTGCCCTCCTGCCTCTGGATTTTAGGTGTTGATTTCTGGAT  
CACGGGATACCACTTCTGTCCACACCCCGACCAGGGGCTAGAAAATTTGTTTGAGATTTTATA  
TCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCC  
TAATTGTTTTACTTGTTAACTTGTTCTTGTCCTTGGGCACTTGGCCTTTGTCTGCTCTCAGTG  
TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTGCACCTCAACGTCGTG  
TGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTACCTTATATCTGTGTTGTTATCCAGGGCTCC  
AGACTTCCTCCTCTGCCTGCCCCACTGCACCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC  
AGCCCAGTCTTGGCTTCTTGTCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG  
CAGAACACCAGGGCCTGGCCCAGTGGATTCATGGTGATCATTAAAAAGAAAAATCGCAACCAA  
AAAAAAAAA

**FIGURE 128**

MARPGMERWRDRLALVTGASGGIGAAVARALVQGLKVVGCAITVGNIEELAAECKSAGYPGTLI  
PYRCDLSNEEDILSMFSAIRSQHSQVDICINNAGLARPDITLLSGSTSGWKDMFNVNVLALSICTR  
EAYQSMKERNVDDGHIININMSGHRVLPVSVTHFYATKYAVTALTEGLRQELREAQTHIRATC  
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

**Important features of the protein:****Signal peptide:**

amino acids 1-17

**N-myristoylation sites.**

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,  
199-205

**Short-chain alcohol dehydrogenase.**

amino acids 30-42, 104-114



**FIGURE 129**

AACTTCTACATGGGCCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT  
CATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTTCAAGTGCAGAGTCTCAGTTGCCCGGGAGC  
ACCTCCCCTCCCAGGCAGTCTGCTCAGAGGGCCTCGGCCAGAATTCAGTTCTGGTTTCATGC  
CAGCCTGTAAAAGGCCATGGAACCTTTGGGTGAATCACCGATGCCATTTAAGAGGGTTTTCTGCCA  
GGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTCAATTCAGTAGCCACCAGCCACCTGTGG  
CCGTTGAGTGCTTGAAATGAGGGAAGTGAAGAAATTAATTTCTCATGTATTTTTCTCATTTATTTA  
TTAATTTTTAACTGATAGTTGTACATATTTGGGGTACATGTGATATTTGGATACATGTATACAA  
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTTATTTTTATTCTTT  
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC  
CTCTGCCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT  
GCACCACAATGCCCAACTAATTTTTGTATTTTTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG  
CTGGCCTTGAACTCCTGGCCTCAAACAATCCACTTGCCTCGGCCTCCCAAAGTGTTATGATTACA  
GGCGTGAGCCACCGTGCTGGCCTAAACATTTATCTTTCTTTGTGTGGGAACCTTGAAATTAT  
ACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT  
ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTCTATCCTTCCC  
AACCTCTGATCACCTCATTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCACATGTG  
AGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCCTG  
TTCCATCCATGTTGCTGCAAATGACAGGATTTGTTCTTAATTTCAATTAATAAATAACACACATG  
GCAAAAA

# **FIGURE 130**

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV  
KGGHTLGESEMPFKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

**Important features of the protein:**

**Signal peptide:**

amino acids 1-18

**N-myristoylation site.**

amino acids 86-92

**Zinc carboxypeptidases, zinc-binding region 2 signature.**

amino acids 68-79

**FIGURE 131**

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC  
 ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCTTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG  
 TCCATCTGTGTGCTGCGTGGGATGCGGGTTTCATTTACTGTAATGATCGCTTTCTGACATCCATTCCAACAG  
 GAATACCAGAGGATGCTACAACCTCTACCTTCAGAACAACCAAATAAATAATGCTGGGATTCCTTCAGAT  
 TTGAAAACTTGCTGAAAGTAGAAAGAATATACCTATACCACAACAGTTTAGATGAATTTCTACCAACCT  
 CCCAAAGTATGTAAGAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCATTATGATTCACTTTCAA  
 AAATTCCTTATCTGGAAGAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA  
 TTCGAGACAGCAACTATCTCCGACTGCTTTTCTGTCCCGTAATCACCTTAGCACAAATTCCTGGGGTTT  
 GCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTTCATCACCATCTCTTCAAG  
 GTCTCACTAGTCTAAAACGCCCTGGTTCTAGATGGAAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTT  
 TTCTTCAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGGGAATTCCTGACTGCTGCACCAGTAAACCT  
 TCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTCCCCCAATGCTTTTT  
 CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTT  
 GATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGAAGATGAAATG  
 GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAA  
 AGGTTTCGTGGGATGGCTATTAAGGATCTCAATGCAGAACTGTTTGATTGTAAGGACAGTGGGATTGTAAGC  
 ACCATTAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCAAGGACAGTGGCCAGCTCCAGTGAC  
 CAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAA  
 CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTCATATCTCTTGAAACTTGCTCTACCTATG  
 ACTGCTTTGAGACTCAGCTGGCTTAACTGGGCCATAGCCCGGCATTTGGATCTATAACAGAAACAATTGT  
 AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTACCCCTATAAAGTATGCATGGTTC  
 CCATGGAAACCAGCAACCTCTACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAAACTGCACCCCTT  
 CGAATGTACAACCCCTACAACCACCCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACC  
 TTTGGCTGCCATCATTGGTGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTGGTATG  
 TTCATAGGAATGGATCGCTCTTCTCAAGGAACGTGCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT  
 GCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAAATCAGGGAACCTCTTTTCAGATGTTACCAAT  
 AAGCAATGAACCCATCTCGAAGGAGGAGTTGTAATACACACCATATTTCTCCTAATGGAATGAATCTGT  
 ACAAAAACAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC  
 TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTTTGGGTTTTTTAAACCTAAGGGAGGTGATG  
 GT



**FIGURE 133**

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCCAGGCCACCCAGGCTTCTTGCCA  
 GCCCTGCCGGGCCACTTGTCTTCATGTCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG  
 TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG  
 GCGGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT  
 CCTCCTCACTGCCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT  
 GCGCTGGAGAGAGCCACCGTCTTCTTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT  
 GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCGGGAGAGTGGGCCCAGGAGCCCCTGC  
 TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC  
 CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCAGCTGACCCTCCAGCCCGG  
 GTTTTGGAAGCTCCACATGCCTGGATCCACACTGATGCCTCCTTGGTGTACCCACGTTCCGGG  
 CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCTGGTGCAGCTGCTGGGAACCGGACG  
 GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC  
 AGGCTACTGCCTGTCCACCAACTGCTCTTCTCCTCTGGGCCAGAATGAGGGGATGCACACAGG  
 GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC  
 AGAGCTGAGGCCATCGGATACGCCTACCTTACCCGGGACATCTTCATGGAAAACATCATGTTCTG  
 TGGAAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA  
 AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA  
 TATCAGCAGCATTTTTTCGAGGAGAGTGAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT  
 TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGC  
 AATTCTCTTGCCTCATCCTCCCGAGTAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAAT  
 TTTTATATTTTTTTAGTAGAGACAGGGTTTCATCATGTGCTCATGCTGGTCTCGAACTCCTGAT  
 CTCAAGAGATCCGCCCACCTCAGGCTCCCAAAGTGTTGGGATTATTAGGTGTGAGCCACCGTGTCTG  
 GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC  
 ATGGGGGCTCTCTCCCTAGATGGCTGCTCCTCCACACACAGCCACAGCAGTGGCAGCCCTGG  
 GTGGCTTCCTATACATCCTGGCAGAATACCCCCAGCAAACAGAGAGCCACACCCATCCACACCG  
 CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCC  
 TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG  
 ATAAGCAAAGCCACCCGACACCCAATCTTGGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG  
 GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAA

## **FIGURE 134**

MSARGRWEGGGRRACRGSGLGLARAQGAERTVSSEQRPMASLGLLLLLLLLTALPPLWSSSLPGLD  
 TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV  
 GMLGEKLEAAIQRSYLHLKLSDPKYLRQFQTLTLPQGFWKLPHAWIHTDASLVYPTFGPQDSFSEE  
 RSDVCLVQLLGTGTDSSSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD  
 YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSWQKQOEGCFG  
 EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSCILIP  
 SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-26

#### **Transmembrane domain:**

amino acids 39-56

#### **Tyrosine kinase phosphorylation sites.**

amino acids 149-156, 274-282

#### **N-myristoylation sites.**

amino acids 10-16, 20-26, 63-69, 208-214

#### **Amidation site.**

amino acids 10-14

#### **Glycoprotein hormones beta chain signature 1.**

amino acids 230-237

# **FIGURE 135**

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCCTCTGTGGGGCTTCTTTCCCGTCCTGCTGCTG  
CTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG  
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGAGTTGTTCCAGGGGTGAAGC  
CTCAGGACTGGATCTCGGCGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTTCCTTAAG  
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGGAAGTTGTATCTCC  
AGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG  
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT  
CCACCTTCTTACTTTATTAAAAGGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCCAATGGT  
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC  
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT  
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAAATCATCTGGCAAATCTAGCAGCGGCAGCAG  
TAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTTCAC  
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGAAAACCGTGTGAAGCAACTACTATAAACTT  
GAGTCATCCCGACGTTGATCTCTTACAACGTGTATGTT  
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGTAT  
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAATTAT  
ATGAACTACTATACATTATGTATATTAATTAACATCTTAATCCAGAAATCAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAA

[illegible]

Important features of the protein:

amino acids 1-23

amino acids 161-182

amino acids 184-187

amino acids 37-40, 236-239

amino acids 151-154

amino acids 33-38, 36-41, 38-44, 229-234

amino acids 238-241

amino acids 229-236





[illegible]

Important features of the protein:

amino acids 1-42

amino acids 192-195, 225-228

amino acids 42-47, 46-51, 136-141

# **FIGURE 139**

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTTCGGCATCCAGCC  
TAGCGTGTCCACGATCGCGGCTGGGCTCCGGGACTTTCGCTACCTGTTGCGTAGCGATCGAGGTGC  
TAGGGATCGCGGTCTTCCCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCCCTCTGCCAGAGCGGAA  
CACGGAGCGGAGCCCCAGCGCCGAACCTCGGCTGGAGCCAGTTCTAACTGGACCACGCTGCC  
ACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGATTTTGTGTTG  
GGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTACCTTGTGGAAAAAGGAGCATCTCACAGT  
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG  
CCTTCCTGGCTTTGTGACGTCATCAGGAACCTCAATTCTCCTGCACTGCTGGAAGACAGTGTGA  
TAAGACAAGCAAAGCAGCTGGAAGAAAGTATAGTCTTTTATGGAGATGAAACCTGGGTTAAATTA  
TTCCCAAAGCATTGTGTGGAATATGATGGAACAACCTCATTTCGTGTGAGATTACACAGAGGT  
GGATAATAATGTCACGAGGCATTGGGATAAGTATTAAGAGAGGAGATTGGGACATATTAATCC  
TCCACTACCTGGGGCTGGACCACATTGGCCACATTTCAGGGCCCAACAGCCCCCTGATTGGGCAG  
AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGA  
GACGCCCTTACCAATTTGCTGGTTCTTTGTGGTGACCATGGCATGTCTGAAACAGGAAGTCACG  
GGGCTCCTCCACCGAGGAGGTGAATACACCTCTGATTTTAATCAGTTCTGCGTTTGAAGGAAA  
CCCGGTGATATCCGACATCCAAAGCACGTCCAATAGACGGATGTGGCTGCGACACTGGCGATAGC  
ACTTGGCTTACCGATTCCAAAGACAGTGTAGGGAGCCTCCTATTCCCAGTTGTGGAAGGAAGAC  
CAATGAGAGAGCAGTTGAGATTTTACATTTGAATACAGTGCAGCTTAGTAAACTGTTGCAAGAG  
AATGTGCCGTCATATGAAAAGATCCTGGGTTTGAGCAGTTTAAATGTGAGAAAGATTGCATGG  
GAACTGGATCAGACTGTACTTGGAGGAAAAGCATTGAGAAGTCCTATTCAACCTGGGCTCCAAGG  
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCAG  
TTCTCACCCCTGCTCCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCA  
CTGTCTCTCCTGGGTTTTCTCTGCTCTTTATTTGGTGATCCTGGTTCTTTCGGCCGTTACAGT  
CATTGTGTGCACCTCAGCTGAAAGTTCGTGCTACTTCTGTGGCCTCTCGTGGCTGGCGGCAGGCT  
GCCCTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGGCAGTGCCTGGAC  
AGGGGGCCTCAGGGAAGGACGTGGAGCAGCCTTATCCCAGGCTCTGGGTGTCCCGACACAGGTG  
TTCACATCTGTGCTGTGAGGTGAGTGCCTCAGTTCTTGGAAGCTAGGTTCTGCGACTGTTAC  
CAAGGTGATTGTAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAA  
TCGGACAGCCTCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCTGGA  
CACTCAGGAGGGTCAAAGGAGACTTGGTGCACCACTCATCTGCCACCCCAAGATGCATCCT  
GCCTCATCAGGTCCAGATTTCTTTCCAAGGCGGACGTTTTCTGTGGAATTCTTAGTCTTGGCC  
TCGGACACCTTCATTCTGTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCAC  
ACTCAGATCCACAGAGCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCCC  
ACCCCAACCTGCACAGCCCTCATCCCTCTTGGCTTGAGCCGTCAGAGGCCCTGTGCTGAGTGT  
CTGACCGAGACACTCACAGCTTTGTATCAGGGCACAGGCTTCTCGGAGCCAGGATGATCTGTG  
CCACGCTTGCACCTCGGGCCCATCTGGGCTCATGCTCTCTCTCCTGCTATTGAATTAGTACCTAG  
CTGCACACAGTATGTAGTTACCAAAAGAATAAACGGCAATAATTGAGAAAAAAA

## **FIGURE 140**

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPPPLF  
 SKVVIVLIDALRDDVFVFGSKGVKMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF  
 VDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPKHFEYDGTTSFFVSDYTEVDNNV  
 TRHLDKVLKRGDWDILILHYLGLDGHIGHISGPN SPLIGQKLEMDSVLMKIHTSLQSKERETPLP  
 NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-34

#### **Transmembrane domain:**

amino acids 58-76

#### **N-glycosylation sites.**

amino acids 56-60, 194-198

#### **N-myristoylation sites.**

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,  
 275-281, 278-284

#### **Amidation site.**

amino acids 154-158

#### **Cell attachment sequence.**

amino acids 205-208

# **FIGURE 141**

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA  
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAACT  
TCCCTATAGAAAACAAC TGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACTTAAAC  
CCGAAGAAATT CAGCATT CATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGAATCTCAT  
AGCAGTTCCAGATAAAAACTACATACGCCCAGAGATCTTCTTTGCATTAGCCTCATCCTTGAGCT  
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGGAGTTTGTCTCTAC  
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT  
GGCTGCCCCAAAAGGAATCAGCACGCCGGCCCTTCATCTTTTATAGGGCTCAGGTGGGCTCCTGGA  
ACATGCTGGAGTCGGCGGCTCACCCCGGATGGTTCATCTGCACCTCCTGCAATTGTAATGAGCCT  
GTTGGGGTGACAGATAAATTTGAGAACAGGAAACACATTGAATTTTCATTTCAACCAAGTTTGCAA  
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAACTGCCCCATTGAACGCCTTCCTCGCTA  
ATTTGAACTAATTGTATAAAAAACACCAAACCTGCTCACT

## **FIGURE 142**

MLLLLEYNFPIENNCQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVLD SGNLI AVDPKNYIRPEI  
FFALASSLSSASA EKGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAAQKESARRPFI  
FYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**  
amino acids 33-36

**N-myristoylation site.**  
amino acids 50-55, 87-92

**Interleukin-1**  
amino acids 37-182



### **FIGURE 144**

MLGLPWKGGLSWALLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY  
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP  
WMTQFSLLNKTCLEGFH

**Important features of the protein:**

**Signal peptide:**

amino acids 1-25

**N-glycosylation sites.**

amino acids 117-121, 139-143

**N-myristoylation site.**

amino acids 9-15





$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}, \quad z = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix}, \quad w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix},$$

Important features of the protein:

amino acids 1-26

## amino acids 157-171

## amino acids 98-102, 110-114

## amino acids 76-83

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

amino acids 62-66

# **FIGURE 147**

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATT  
 CCTGATGATTTATAGACTCAAAGAAACTCATGTTTCAGAAGCTCTCTTCTTCTGGCCTCCTCT  
 CTGTCTTCTTTCCCTCTTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG  
 AAATCTTTCATTTTGCTTGTCAGTGGGGTAGGTCAGTCTTAGTTTTTATTTTTTGAAATTT  
 CAACTTTCAGATTCAGGGGGTACATGTGAAGGTTTGTATGAGTATATTGCATGATGCTGAGG  
 TTTGGGGT

## **FIGURE 148**

MFRSSLLFWPPLCLLSLFLILISSIYSESKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE  
GLFYEYIA

**Important features of the protein:**

**Signal peptide:**

amino acids 1-25

**N-myristoylation site.**

amino acids 62-68

**FIGURE 149**

GTCTCCGCGTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGACTTGAC  
 TCCCGCGCGCCCCAACCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGCCGCCAGTCC  
 CGGCCCCCTCTCCCGCCCCACCCACCTCCTGGCTCTTCCTGTTTTTACTCCTCCTTTTCATTGATA  
 ACAAAAGCTACAGCTCCAGGAGCCCAGCGCCGGGCTGTGACCCAAGCCGAGCGTGGAAGAAATGGGGTT  
 CCTCGGGACCGGCACTTGGATTCTGGTGTTAGTGCTCCCGATTCAAGCTTTCCCAAACCTGGAGGAA  
 GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAA  
 GCAGAAGAAGACAAGATTAAAAAACATATCCTCCAGAAAACAAGCCAGGTGAGAGCAACTATTCTTT  
 TGTTGATAAATTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTA  
 TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACCTG  
 ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAAATTTCAAGATGATCCAGATGGTCTTCA  
 TCAACTAGACGGGACTCCTTTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTATGAAG  
 AAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAAACTACTTAATCTCGGCCTTATCACAGAAAGC  
 CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAAATTAATCTCAAAGGAAGCCAACAA  
 TTATGAGGAGGATCCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG  
 TGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA  
 TTAACCTTGACAAATGGCTTGGAAGGAGAACTAAAACCTACAGTGAAGACAACCTTGAGGAACTCCA  
 ATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAAGCAAAAGAGAAAGAAA  
 CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAAATATGGAACAATATCT  
 CCAGAAGAAGGTGTTTCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAACAAGCT  
 AGAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCCAGCACCATCAGAGAAGAGTCATGAAGAAA  
 CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAA  
 GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGC  
 CATCAGAAAAAATATGAATGGTTGAAGAAACATGACAAAAGGGAAATAAAGAAGATTATGACCTTT  
 CAAAGATGAGAGACTTCATCAATAAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA  
 GAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAATAAATGGCAAAAGATCCAGGAGTCTTTCAA  
 CTGTTTCAGAAAACATAATATAGCTTAAAACACTTCTAATTCTGTGATTAAAAATTTTTGACCCAAGG  
 GTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGCTTTCTTCCC  
 GTAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

**FIGURE 150**

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG  
 QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK  
 FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE  
 VLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVTMAAIQDGLAKGENDETVSNTLTLTNGLE  
 RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV  
 SYLENLDEMIALQTKNKLEKNATDNISKLFAPSEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD  
 NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHKKGKEDYDLSKMRDFINKQADAYVEKGILDK  
 EEAEAIKRIYSSL

**N-glycosylation sites:**

amino acids 68-71, 346-349, 350-353

**Casein kinase II phosphorylation site:**

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-  
 220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-  
 341, 369-372, 382-385, 386-389, 387-390

**N-myristoylation sites:**

amino acids 143-148, 239-244

## FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCTGTGGAGCTCAA  
CATGGTCCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC  
ATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCAATTAAAGGTGAAGAGATCAGC  
GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCCGTCATCCTGGGTGTCCAGGGTGAAG  
CCAGTGCCTGTGCATGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG  
AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTCACCTTCTACCGCGGGACATGGGGCTCACC  
TCCAGCTTCGAGTTCGGCTGCCTACCCGGGCTGGTTCCCTGTGCACGGTGCCTGAAGCCGATCAGCC  
TGTCAGACTCACCCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCATCACAGACTTCTACTTCC  
AGCAGTGTGACTTAGGGCAACGTGCCCCCAGAACTCCCTGGGCAGAGCCAGCTCGGGTGAGGGGT  
GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG  
TGGGCACCTGACCACTTTTGTCTTCTGGTTCCCAAGTTTGATAAATTCTGAGATTTGAGCTCAGT  
CCACGGTCTCCCCACTGGATGGTGTCTACTGCTGTGGAACCTTGTAACCAATGATGGGGTAA  
CTGGGAATAACATGAAAAGATTTCTGTGGGGGTGGGGTGGGGGAGTGGTGGGAATCATTCCTGCT  
TAATGGTAACGTACAAGTGTACCCCTGAGCCCCGAGGCCAACCCATCCCCAGTTGAGCCTTATA  
GGGTGAGTAGCTCTCCACATGAAGTCCTGTCACTCACCCTGTGCAGGAGAGGGAGGTGGTCTATA  
GAGTCAGGGATCTATGGCCCTTGGCCAGCCCCACCCCTTCCCTTTAATCCTGCCACTGTCTATA  
TGCTACCTTTCTATCTCTTCCCTCATCATCTTGTGTGGGCATGAGGAGGTGGTGTGTGAGAA  
GAAATGGTTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTCTTTTAAAAACCCAA  
GATACAATCAAAATCCCAGATGCTGGTCTCTATTCCCATGAAAAAGTGTCTGATGACATATTGAGA  
AGACCTACTTACAAAGTGGCATATATTGCAATTTATTTTAATTAAGATACCTATTTATATATT  
TCTTTATAGAAAAAGTCTGGAAGAGTTTACTTCAATTGTAGCAATGTCAGGGTGGTGGCAGTAT  
AGGTGATTTTTCTTTTAATCTGTGTAATTTATCTGTATTTCCTAATTTTCTACAATGAAGATGA  
ATTCCTTGTATAAAAAATAAGAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGA  
TTGTCTCAGCCTCCACTTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTTGG  
TTGTAGTAGTGATCAGGAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGT  
GTGGCTGGAATCTCTGGGTAAAGAACTTAAAGAACAAAAATCATCTGGTAATCTTTCCCTAGAAG  
GATCACAGCCCTGGGATTTCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTTGA  
ATTGTGTCCCCCTCAAATTCACATCCTTCTTGGAATCTCAGTCTGTGAGTTTATTTGGAGATAAG  
GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCTGCTGGATGAAGGTAGACCTAAATTCATAT  
GACTGGTTTCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA  
AAGATGAAGGCAGAGATCGGAGTTTTCAGCCACAAGCTAAGAAACACCAAGGATTGTGGCAACC  
ATCAGAAGCTTGGAAGAGGCAAAGAAGAAATCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCTG  
CTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAAGATAAATTTCCGGCTGTTTTAA  
GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAATGATCCCTGT  
CTCCTCGTGTTTACATTCTGTGTGTGTCCCTCCCACAATGTACCAAAGTTGTCTTTGTGACCAA  
TAGAATATGGCAGAAGTGATGGCATGCCACTTCCAAGATTAGGTTATAAAAGACACTGCAGCTTC  
TACTTGAGCCCTCTCTCTCTGCCACCCACCGCCCAATCTATCTTGGCTCACTCGCTCTGGGGG  
AAGCTAGCTGCCATGCTATGAGCAGGCCATATAAGAGACTTACGTGGTAAAAATGAAGTCTCCT  
GCCACAGCCACATTAGTGAACTTAGAAGCAGAGACTCTGTGAGATAATCGATGTTTGTGTTTT  
AAGTTGCTCAGTTTTGGTCTAACTTGTATGCAGCAATAGATAAATAATATGCAGAGAAAGAG

## **FIGURE 152**

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEESISVVPNRWLDASLSPVILGVQGG  
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP  
VRLTQLPENGGWNAPITDFYFQQCD

### **N-myristoylation sites.**

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

### **Interleukin-1 signature.**

amino acids 111-131

### **Interleukin-1 proteins.**

amino acids 8-29, 83-120, 95-134, 64-103



# **FIGURE 153**

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAAGTTGCTCGAGTTAGAATTGTCTGCAATGCCCCGC  
CCTGCAGAAATCTGTGAGCTCTTTCCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTCTTGG  
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCAGTGCAGGCTTGACAAGTCCAAC  
TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA  
CAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT  
ATCTGATGAAGCAGGTGCTGAACCTTCACCCCTTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTTT  
CAGCCTTATATGCAGGAGGTGGTGGCCTTCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA  
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC  
TTGGAGAGAGTGAGAGATCAAAGCAATTGGAGAACTGGATTTGCTGTTTATGTCTCTGAGAAAT  
GCCTGCATTGACACAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTTCCCTGCTAGAAATAA  
CAATTAGATGCCCCAAAGCGATTTTTTTTAAACAAAAGGAAGATGGGAAGCCAACTCCATCATG  
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTTGTTTATA  
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTTTCATTGTAAGTGGTGTTC  
TATACACAGAAAAACAATTTATTTTTTAAATAATTGTCTTTTCCATAAAAAAGATTACTTTCCAT  
TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAA  
TGTATTTATTATTATTATAAGACTGCATTTTATTTATATCATTTTATTAATATGGATTTATTTAT  
AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT  
AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGGATATCCC

**FIGURE 154**

MAALQKSVSSFLMGTLATSCLLLLALLVQGGAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL  
ADNNTDVRLIGEKLFGVSMSERCYLMKQVLNFTLEEVLFPQSDRFQPYMQEVVPFLARLSNRLS  
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

**Important features of the protein:****Signal peptide:**

amino acids 1-33

**N-glycosylation sites.**

amino acids 54-58, 68-72, 97-101

**N-myristoylation sites.**

amino acids 14-20, 82-88

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 10-21

## FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT  
CAGTCAGTGCCCGACTTGTGACTGAGTGTGAGTGCCCGAGCATGTACCAGGTCAGTGAGAGGGC  
TGCTTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC  
TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGGGTGCAGAT  
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCTACAGGTGGTTGCAT  
TCTTGGAATGGTCATGGGAACCCACACCTACAGCCACTGGCCCAGCTGCTGCCCCAGCAAAGGG  
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCCTAGAGCCTGCTAG  
GCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGCCATCT  
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAGGACCTGTACCACGCCCCGT  
TGCTGTGCCCCGCACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCGGGGCAACTCGGA  
GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCAGCCACA  
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTTCTTAGCTTGTGTGTGTGTGCGGCCCCGT  
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTTGGGAAACCTGGAGCCAGGTGTACA  
ACCACTTGCCATGAAGGGCCAGGATGCCCAGATGCTTGGCCCCGTGTGAAGTGCTGTCTGGAGCAG  
CAGGATCCCGGGACAGGATGGGGGGCTTTGGGGAAAACCTGCACTTCTGCACATTTTGAAAAGAG  
CAGCTGCTGCTTAGGGCCGCGGAAGCTGGTGTCTGTCTATTTCTCTCAGGAAAGGTTTTCAAA  
GTTCTGCCCATTTCTGGAGGCCACCACTCCTGTCTCTTCTCTTTTCCCATCCCCTGCTACCCGT  
GCCCAGCACAGGCACTTTCTAGATATTTCCCCCTTGCTGGAGAAGAAAGAGCCCCCTGGTTTTATT  
TGTTTGTTTACTCATCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTT  
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTTATCCAAATAAATAT  
CTTTATTTAAAAATGAAAAA

[illegible]

Important features of the protein:

amino acids 1-32

amino acids 136-140

amino acids 127-135

amino acids 44-50, 150-156

# **FIGURE 157**

CCGGCGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC  
GACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC  
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACAGTGTGCAACAGGGGACTATTCA  
ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTTGAAGGCCACCAA  
GATTTGTGTGACGGGCAAAAGCAACTTCCAGTCCTACAGCTGTGTGAGGTGCAATTACACAGAGG  
CCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCCTACATCGGCTTCCCTGTA  
GAGCTGAACACAGTCTATTTTCATTGGGGCCATAATATTCCTAATGCAAATATGAATGAAGATGG  
CCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAA  
AGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTGAAGAAGAATGAGGAGACA  
GTAGAAGTGAACCTTACAACCACTCCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC  
TATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA  
TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCCTACTTGTGGC  
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCCACAAACAGGCGTCCCTTTCCCTCT  
GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGCTGTCTCTGCTGGTGGCCA  
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT  
TCTACCACCACACTACTGCCCCCATTAAGGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCA  
TCACACAATTTGTTACTTCACTGAATTTCTTCAAAACCATTGCAGAAGTGAGGTTCATCCTTGAAA  
AGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAAGAAGGCA  
GCAGACAAAGTCGTCTTCCTTCTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA  
GAGCGAGGGCAGTCCCAGTGAGAACTCTCAAGACCTCTTCCCCCTTGCCTTAACCTTTTCTGCA  
GTGATCTAAGAAGCCAGATTCTGACACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA  
AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT  
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTGAGCAGGAAAAAGATCACAAGCCTGCCACG  
ATGGCTGCTGCTCCTTGTAG

## FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLPGLDLRLRVEPVTTSVATGDYSILMNVS  
 LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP  
 NANMNEDGPSMSVNFTSPGCLDHIMYKKKCVKAGSLWDPNITACKKNEETVEVNFTTTPLGNRYMALIQH  
 STIIGFSQVFEPHQKKQTRASVVIPTVGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPTGVPFPLDNNK  
 SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFL  
 QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVELLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA  
 ENLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVC PKYHLMKDATAFCAELLHVKKQVSAGKRSQACHD  
 GCCSL

### Important features of the protein:

#### Signal peptide:

amino acids 1-14

#### Transmembrane domain:

amino acids 290-309

#### N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283  
 - 287

#### cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

#### Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

#### N-myristoylation site.

amino acids 116-122

#### Amidation site.

amino acids 488-452



## **FIGURE 160**

MTVKTLHGPMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN  
QRVMSRNIERSRSTSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQKEDISMNSVPIQQETLVV  
RRKHQGCSVSFQLEKVLVTVGCTCVTPVIHHVQ

**Signal sequence:**

amino acids 1-30

**N-glycosylation site.**

amino acids 83-87

**N-myristoylation sites.**

amino acids 106-111, 136-141



**FIGURE 161**

ACACCTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCAGG  
 ACAGAGAGTGCACAACTACCCAGCACAGCCCCCTCCGCCCCCTCTGGAGGCTGAAGAGGGATTTC  
 CAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCCCTTGGGGGGGGGAGCAGC  
 AGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCTGTGCCCTGGTTCTGTGTCTCT  
 TGGCACTGGGGCCGAAGCCAGTGGTCTTTCTCTGGAGAGGCTTGTGGGGCCTCAGGACGCTACC  
 CACTGCTCTCCGGGCTCTCCTGCCGCTCTGGGACAGTGACATACTCTGCTGCTGCTGGGACAT  
 CGTGCTGCTCCGGGCCCCGTGCTGGCGCCTACGCACCTGCAGACAGAGCTGGTGCTGAGGTGCC  
 AGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGCCGTGCATGGGCACTGG  
 GAAGAGCCTGAAGATGAGGAAAAGTTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCCTAGGAA  
 TGCCTCTCTCCAGGCCAAGTCTGTCTCTCCTTCCAGGCCTACCCTACTGCCCGCTGCGTCTCTGC  
 TGGAGGTGCAAGTGCCTGCTGCCCTTGTGCAGTTTGGTCAGTCTGTGGGCTCTGTGGTATATGAC  
 TGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTCTTACTCAGCCCAGGTACGAGAA  
 GGAACCTCAACACACACAGCAGCTGCCCTGCCCTGGCTCAACGTGTGACGAGATGGTGACA  
 ACGTGATCTGGTTCTGAATGTCTCTGAGGAGCAGCACTTCGGCCTCTCCCTGTACTGGAATCAG  
 GTCCAGGGCCCCCAAAACCCCGGTGGCACAAAAACCTGACTGGACCGCAGATCATACCTTGAA  
 CCACACAGACCTGGTTCCCTGCCTCTGTATTTCAGGTGTGGCCTCTGGAACCTGACTCCGTTAGGA  
 CGAACATCTGCCCTTTCAGGGAGGACCCCGCGCACACCAGAACCTCTGGCAAGCCGCCGACTG  
 CGACTGCTGACCTGCAGAGCTGGCTGCTGGACGCACCGTGTCTGCTGCCCGCAGAAGCGGCACT  
 GTGCTGGCGGGCTCCGGGTGGGGACCCCTGCCAGCCACTGGTCCCACCGCTTTCTGGGAGAACG  
 TCACTGTGGACAAGTTCTCGAGTTCCCATGCTGAAAGGCCACCCCTAACCTCTGTGTTCAGGTG  
 AACAGCTCGGAGAAGCTGCAGCTGCAGGAGTGCTTGTGGGCTGACTCCCTGGGGCTCTCAAAGA  
 CGATGTGCTACTGTTGGAGACACGAGGCCCCAGGACAACAGATCCCTCTGTGCCCTTGAACCCA  
 GTGGCTGTACTTCACTACCCAGCAAAGCCTCCACGAGGGCAGCTCGCCTTGGAGAGTACTTACTA  
 CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCCTG  
 CCCCATGGACAAAATACATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCCTACTCTTTGCCG  
 CTGCGCTTTCCCTCATCTCTCTCTCAAAAAGGATCACGCGAAAGGGTGGCTGAGGCTCTTGAAA  
 CAGGACGTCCGCTCGGGGGCGGCCAGGGGCCGCGGGCTCTGCTCCTCTACTCAGCCGATGA  
 CTCGGGTTTCGAGCGCCTGGTGGGCGCCCTGGCGTCCGGCCTGTGCCAGCTGCCGCTGCGCGTGG  
 CCGTAGACCTGTGGAGCCGTGCTGAACTGAGCGCGCAGGGGCCCGTGGCTTGGTTTCACGCGCAG  
 CGGCGCCAGACCTGCAGGAGGGCGGCGTGGTGGTCTTGTCTTCTCTCCCGGTGCGGTGGCGCT  
 GTGCAGCGAGTGGCTACAGGATGGGGTGTCCGGGCCCGGGGCGCACGGCCCGCACGACGCTTCC  
 GCGCCTCGCTCAGCTGCGTGCTGCCCGACTTCTTGCAGGGCCGGGCGCCCGGCAGCTACGTGGGG  
 GCCTGCTTCGACAGGCTGCTCCACCCGACGCGTACCCGCCCTTTCCGACCGTGGCCGCTCTT  
 CACACTGCCCTCCCACTGCCAGACTTCTTGGGGGCCCTGCAGCAGCCTCGCGCCCCGCTTCCG  
 GCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCCGGGCCCTTCAGCCAGCCCTGGATAGCTACTTC  
 CATCCCCGGGGACTCCCGCGCCGGGACGCGGGTGGGACCAGGGGGGGACCTGGGGCGGGGA  
 CGGGACTTAAATAAAGGCAGACGCTGTTTTCTAAAAAA

**FIGURE 162**

MPVPWFLLSLALGRSPVVLSSLERLVGFQDATHCSPGLSCLWSDILCLPGDIVPAPGPVLIAPTHLQTELV  
 LRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVLSTFQAYPTARCVLLEV  
 QVPAALVQFGQSVGSSVYDCFEAALGSEVRIWSYTPRYEKELNHTQQLPALPWLNVSAADGNVHLVLNV  
 EEQHFGLSLYWNQVQGPPKPRWHKNLTGPGIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQN  
 LWQAARLRLLTLQSWLLDAPCSLPAAEALCWRAPGGDPCQPLVPPLSWENVTVDKVLEFPLKKGHPNLCVQ  
 VNSSEKLQEQECLWADSLGPLKDDVLLLETRGPDNRSLEPSGCTSLPSKASTRAARLGEYLLQDLQS  
 GQCLQLWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSGAAARG  
 RAALLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAVFHAQRRQTLQEGGVVLLFSP  
 GAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLHPDAVPALFRTVPVFT  
 LPSQLPDFLQPRAPRSGRLQERAEQVSRLQPALDSYFHPGTPAPGRGVGPGAGPGAGDGT

**Signal sequence:**

amino acids 1-20

**Transmembrane domain.**

amino acids 453-475

**N-glycosylation sites.**

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,  
 334-337, 357-360, 391-394

**Glycosaminoglycan attachment site.**

amino acids 583-586

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 552-555

**N-myristoylation sites.**

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,  
 692-697, 696-701, 700-705

**FIGURE 163**

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT  
GCTCACGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCAGTCCAGCAACTTTGA  
AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA  
AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTCAGCGGATCACCCGGAAGTCTGC  
AACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT  
GTCAGTGGGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTGAGTCTCTGACGACACTAC  
CCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTGAGATGATTGTTTCATCCTA  
CCCCACGCCAATCCGTGCAGGCGATGGCCACCGGCTAACCTGGAGACATCTTCCATGACCTG  
TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGAGGGAAGCAGAGAGA  
ATATGAGTCTTCGGCCTGACCCCTGACACAGAGTTCCTTGGCACCATCATGATTGCGTTCCCA  
CCTGGGCCAAGGAGAGTGGCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC  
TACTCCTTCTCCGGAGCCTTCTGTCTCCATGGGCTTCTCGTCGAGTACTCTGCTACCTGAG  
CTACAGATATGTACCAAGCCGCTGCACCTCCCAACTCCCTGAACGTCCAGCGAGTCTGACTT  
TCCAGCCGCTGCGCTTCTATCCAGGAGCACGTCTGATCCCTGTCTTTGACCTCAGCGGCCACG  
AGTCTGGCCACGCTGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGCAGGAGC  
TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC  
CCTCCAACGTGCCACCTCCCGAGATCCTTCCCCACTGTCTATGCCCAAACGCTGCCCTTGAG  
GTCGGGCCCCCATCTATGCACCTCAGGTGACCCCCGAAGCTCAATTCCCATTTACGCCCCACA  
GGCCATCTTAAGGTCCAGCCTTCTCCTATGCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT  
CCTATGGGGTATGCATGGAAGGTTCTGGCAAAGACTCCCCACTGGGACACTTTCTAGTCTTAA  
CACCTTAGGCCTAAAGGTGAGCTTTCAGAAAGAGCCACCAGCTGGAAGCTGCATGTTAGGTGGCCT  
TTCTCTGCAGGAGGTGACCTCCTTGGCTATGGAGGAATCCCAAGAAGCAAAATCATGACCAGC  
CCCTGGGGATTGACACAGACAGAATCTGACCCAAATGTGCTACACAGTGGGGAGGAAGGACAC  
CCACAGTACCTAAAGGGCCAGCTCCCCCTCCTCTCCTCAGTCCAGATCGAGGGCCACCCCATGTC  
CCTCCCTTTGCAACCTCCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC  
TGCTGGAGTCCCTTGTGTGTCCCAAGGATGAAGCCAAGAGCCCGAGCCCTGAGACCTCAGACCTG  
GAGCAGCCACAGAACTGGATTCTCTTTTCAGAGGCCTGGCCCTGACTGTGAGTGGGAGTCTTG  
AGGGGAATGGGAAAGGCTTGGTGCTTCCCTGTCCTTACCCAGTGTACATCCTTGGCTGTCA  
ATCCCATGCCCTGCCCATGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGC  
AGAGGGAGTGGCATGCAGGGCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATA  
AGGACTGCAGCGGGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGGTGCTCGCTGAGCCCTG  
CAAGGCAGAAATGACAGTGAAGGAGGAAATGCAGGGAAACTCCCGAGTCCAGAGCCCCACCTC  
CTAACACCATGGATTCAAAGTGCTCAGGGAATTTGCCTCTCCTTGCCCCATTCTTGCCAGTTTC  
ACAACTAGCTCGACAGAGCATGAGGCCCTGCCTCTTCTGTCATTGTTCAAAGGTGGGAAGAGA  
GCCTGGAAAAGAACAGGCCTGGAAAAGAACAGAGGAGCTGGGCAGAACAGAACAACTGC  
ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCCTGCAGTCA  
TTCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTGAGCTTCATTCTCTGATAGAACAAAGC  
GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTCTGCAGGCAGGAGTTTCAGACCT  
ATCCTGAGAATGGGTTTGAAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACACAC  
TGTAAGTATGTACAACCTTTGCAAGCTCTGCCTTGGGTTCAAGCCATCTGGGCTCAAATTCAGC  
CTCACCACTCACAGCTGTGTGACTTCAAACAATGAAATCAGTGCCAGAACCTCGGTTCCCTC  
ATCTGTAATGTGGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG  
TCTTTAAAGTGCTTAATAGTGCTGGTACATGGGCAGTGCCCAATAACGGTAGCTATTTAAAAA  
AAAAAAA

## FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTPVYSIEYKTYGERDW  
 VAKKGCQRITRKSCNLTVETGNLTELYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCIS  
 KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHGGKQREYEFFGLTPDT  
 EFLGTIMICVPTWAKESAPYMCRVKTLDPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP  
 PNSLNVQRVLTQPLRFIQEHVLI PVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT  
 YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSSY  
 APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPGQLQKEPPAGSCMLGGLSLQEVTSLAM  
 EESQEAQSLHQPLGICTDRTSDPNVLHSGEEGTPQYLKGQLPLLSSVQIEGHPMSLPLQPPSGPC  
 SPSPDQGPSPWGLLES LVC PKDEAKSPAPETSDLEQPTELDSLFRGLALTVQWES

### Signal sequence.

amino acids 1-17

### Transmembrane domain.

amino acids 233-250

### N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

### N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

## FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAAGTCACCCGGGCCCCGCGGTGGCCACAACATGG  
CTGCGGCGCCGGGGCTGCTCTTCTGGCTGTTTCGTGCTGGGGGCGCTCTGGTGGGTCCCGGGCCAG  
TCGGATCTCAGCCACGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT  
GTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGTGAATTTTA  
AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAACCTTGGGCTGGA  
AGTGTGTAACACAGTTTTTGGATATTTTCCAAAAGATTTGATCAAGGTACTTCATAAATACACGGA  
AGAAGAGCTACATATTCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGGAGGAAGAGATGATT  
TTAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAACTGGAGGACTCTGTACCTGAAGAG  
TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA  
ACTTGACCCTGTGCCTGAGCCCCGAGGCATTTCAGAGCTGATTTCAGAGGATGGAGAAGGTGCTTTCT  
CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTCACACCAGCGGT  
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACACTTTTGAAGAAATTCGCACGATAA  
ATTGAAAGTGCCGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA  
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGC GTTATT  
CATTACAGCAAAGGATTTTCGTTGGCATCAAAATCTAAGTTTGTGTTTACAAAGATTGTTTTTAGTA  
CTAAGCTGCCTTGGCAGTTTGCATTTTTGAGCCAAACAAAATATATTATTTTCCCTTCTAAGTA  
AAAAAAAAAAAAAAAAAAAA

## **FIGURE 166**

MAAAPGLLFWLFLVGLALWWVPGQSDLGRRFSDLKVCGDEECMSMLMYRGKALEDTGPDCRFVN  
 FKKGDDVYVYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEELHIPADETDFVCFEGGRD  
 DFNSYNVEELLGSLELEDSVPEESKKAEEVSQHREKSPEESRGRELDPVPEPEAFRAEDGEDGA  
 FSESTGLQGQPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER  
 EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNLSLFYKDCF

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-22

#### **N-glycosylation site.**

amino acids 294-298

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 30-34

#### **Tyrosine kinase phosphorylation site.**

amino acids 67-76

#### **N-myristoylation sites.**

amino acids 205-211, 225-231, 277-283

#### **Amidation site.**

amino acids 28-32

**FIGURE 167**

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCTAGAGGCCGGGGAAGAGAAGCAAAGCGC  
 AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCCTTAGGACATACACGGGACCCCTAACTTC  
 AGTCCCCCAAACGCGCACCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCGCGGCACAGG  
 CGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTCGGGCAGCTGGGCTCGGGC  
 GGCAGGAGTAGGGCCCCGCGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTCCGCGGGCTGCGCCCTG  
 GGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGCCGCGTGG  
 TCTCGCTGCTGCTGGGCGCCGCGCTGCTGCTGCCGCCACGGAGCCTTCTGCCGCCGCGTGGTCAAGC  
 GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAAGT  
 GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCCTGGCTTGTGAGAGTGAGGGAGGAGTCCCTCCTCA  
 GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG  
 ACAGGGATTCTGATGGTGATTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAACATCTGG  
 TGCCTGCCCAGATCTCTACCAGTGGTCTGATGGAAGCAATTCACAGTACCGAACTGGTACACAG  
 ATGAACCTTCTGCGGAAGTGAAAAGTGTGTTGTGATGTATCACCACCAACTGCCAATCCTGGC  
 CTTGGGGGTCCCTACCTTTACCAGTGGAAATGATGACAGGTGTAACATGAAGCACAAATATATTTG  
 CAAGTATGAACCAGAGATTAATCCAACAGCCCCGTGAGAAAAGCCTTATCTTACAAATCAACCAG  
 GAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGGTATAATTCCCAATCTAATTTATGTTGTT  
 ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTTTGGAACCTGTTGTTTCCAGATGCT  
 GCATAAAAAGTAAAGGAAGAACAACAACTAGTCCAAACAGTCTACACTGTGGATTTCAAAGAGTA  
 CCAGAAAAGAAAGTGGCATGGAAGTATAATAACTCATTGACTTGGTTCCAGAATTTTGTAATCT  
 GGATCTGTATAAGGAATGGCATCAGAACATAGCTTGAATGGCTTGAAATCACAAAGGATCTGC  
 AAGATGAAGTGAAGTCCCTTTGAGGCAATATTAAGTAATTTTATATGTCTATTATTTCA  
 TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAA  
 ACTTCAAACCTCAAGCAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA  
 TGTGTGTTAGAAGCAATTCCTTTTATTTCTTTCACCTTTCATAAGTTGTTATCTAGTCAATGTAA  
 TGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGATAAAA  
 ATGAAGTGTCTAATATTTATTTTATGGCATCTCATTTTCAATACATGCTCTTTTGATTAAAG  
 AAACCTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGT  
 TTTCTCGAAATAATTCATCTTTCAGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGA  
 AATAAGAAGCTATTTTCATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGAT  
 TGTCTAATTTCAATTTGTGCAAGACATGTGCCCTTATAATTTATTTTGTAGCTTAAATTAACAGATT  
 TTGTAATAATGTAACCTTGTAAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAAG  
 TGACATACACAATATAAATCATATGTCTTCACACGTTGCCCTATATAATGAGAAGCAGCTCTCTGA  
 GGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTCCGGGTTTGGG  
 ATTGACACTGGAGGCAGATAGTTGCAAGTTAGTCTAAGGTTTCCCTAGCTGTATTAGCCTCTG  
 ACTATATTAGTATACAAAGAGGTGATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG  
 ACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAAGTACGAAATCGTGTGAAAATGGGTTGG  
 AACCATCAGTGATCGCATATTCAATTGATGAGGGTTTGTCTGAGATAGAAAATGGTGGCTCCTTT  
 CTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCCCTTGTCTTCTCAAGAGAAAGTTGTAAGTCT  
 CTGGTCTTCATATGTCCCTGTGCTCCTTTTAAACCAATAAAGAGTTCTTGTCTTCTGGGGGAAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 168**

MSRVVSLLLGAALLCGHGAFCRRVVSQKVCFADFKHPCYKMAFHELSSRVSFQEARLACESE  
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ  
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWDDRCNMKHNYICKYEPEINPTAPVEK  
PYLTNQPGDTHQNVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSCKGRKTSPNQ  
STLWISKSTRKESGMEV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-21

#### **Transmembrane domain:**

amino acids 214-235

#### **N-glycosylation sites.**

amino acids 86-89, 255-258

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 266-269

#### **N-myristoylation sites.**

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217